In or Out of Control:
Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices,
with Applications in Aviation and Energy Management Domains

Dilhan J. Thilakarathne
Thesis Reading Committee:

prof.dr. Richard P. Cooper  
Department of Psychological Sciences, Birkbeck, University of London, UK

prof.dr. Patricia Lago  
Department of Computer Science, VU University Amsterdam, Netherlands

prof.dr. Max M. Louwerse  
Department of Communication and Information Sciences, Tilburg School of Humanities, Netherlands

prof.dr. John-Jules Ch. Meyer  
Department of Information and Computing Sciences, University of Utrecht, Netherlands

dr. Alexei V. Samsonovich  
Krasnow Institute for Advanced Study, George Mason University, USA

dr. Alexei Sharpanskykh  
Air Transport and Operations, Faculty of Aerospace Engineering, Delft University of Technology, Netherlands

ISBN: 978-94-6332-022-1


The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

Copyright © 2016 by Dilhan J. Thilakarathne.

All rights reserved.

Cover artwork by Angi Pradil and ‘Rubiks Brain’ image by Jason Freeny.
In or Out of Control:

Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices,

with Applications in Aviation and Energy Management Domains

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Exacte Wetenschappen
op woensdag 15 juni 2016 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Dilhan Jayantha Thilakarathne

geboren te Colombo, Sri Lanka
promotor: prof.dr. J. Treur

copromotor: dr. T. Bosse
For My Family
Acknowledgments

It is with great pleasure that I take this opportunity to extend my gratitude towards everyone who helped me to successfully complete this research and thesis. As I am not a self-made man, I cannot forget those who have helped me to get where I am today.

I would like to express my sincere gratitude to my promotor prof.dr. Jan Treur for his continuous support, patience, motivation, availability, and immense knowledge. He always allowed me to realize my ideas and directed me towards accomplishing those with interesting results. There was no stress when working with Jan, and that motivated me to always work harder. I cannot imagine having a better promotor and mentor. Thank you very much Jan for everything!

My PhD has been an amazing (research + travel) experience and I thank Jan again wholeheartedly, for giving me so many wonderful opportunities to attend conferences in addition to his tremendous research support. Getting an opportunity to travel to Warsaw, Prague, and Stockholm for four conferences just immediately after the other was something that cannot be forgotten.

My co-promotor, dr. Tibor Bosse, is a very kind, energetic and helpful person. Research we did for Eurocontrol was a nice challenge in my PhD career. Tibor was always there to help meet all the deadlines of Eurocontrol, and I was able to successfully complete the research and secure two conference papers and one journal paper within one year. Without you, it would not have been a reality. Thank you Tibor for all your support!

I am grateful to all co-authors of research papers I wrote: Amin, Laila, Michel, Jan, and Wim.

I am also thankful to all the members of Agent Systems Research group (now called as Behavioural Informatics group): Adnan, Alexei, Altaf, Amin, Arlette, Charlotte, Daniel, Eric, Jan, Jeroen, Julienka, Lenin, Michel, Natalie, Nataliya, Robbert-Jan, Tibor, and Vera for their nice discussions, meetings, and support; our secretaries: Elly and Mojca for all their support and members of Ambient Assisted Living group: Arjen, Lisette, Peter, Ward, and Wim. Additionally, support I received from Jeroen, Julienka, Lisette, Ward, and Wim for Dutch to English translations, advice on government/accommodation/university related questions/situations, and special advises cannot be forgotten.

I owe my gratitude to my former supervisors: prof.dr. Asoka Karunannada and dr. Ruwan Weerasinghe, and my lecturer for Philosophy of Science, prof.dr. Prian Dias. The experience and exposure I got from them motivated me to start a PhD.

Most importantly, I cannot forget my parents, brother and sister. They always gave me an enormous support, and the faith they had on me were of great strength to me all throughout my life.

I also like to give special thanks to Angi for all the encouragements and nice talks we had. Finally I would like to thank all my dear friends: Dhanushka, Indika, Mahesh, Manoj, Muditha, Prabath, Prashan, Uditha, Viraj, and Waruni for their friendship and support.

Life flows continuously with changes in every moment. I think I had a nice time for the past three years and four months with lot of experience, knowledge, exposure, and opportunities. Nevertheless, for everything there is an end and I think this is the end for my current PhD. I hope what I collected so far will be helpful for me to face the coming changes in my life.

Thank You All!

Dilhan J. Thilakarathne

VU University Amsterdam

20 April 2016
Abstract

Although human cognitive processes may be complex, it is an interesting challenge to achieve some understanding of them. Knowledge of these processes contributes to the development of various application domains that involve human intervention; for example, health, aviation, energy, and safety critical domains. Such domains can only be developed when having sufficient knowledge of human cognitive processes. Understanding human cognition, designing working models based on such understanding, and using these models in application domains are non-trivial research challenges.

Supported by the developments in brain imaging and recording techniques, neurocognitive researchers are discovering more and more knowledge about many cognitive processes. Most of this research focuses on highly important but relatively specialised questions. Therefore, it is difficult to find detailed explanations of a certain cognitive phenomenon as an integrated, coherent process. Nevertheless, having such detailed understanding of cognitive processes is important for various application domains. Understanding the processes in the human brain demands a multidisciplinary approach. Cognitive modelling has increasingly become a prevalent multidisciplinary research theme for this. Dynamic modelling approaches provide insight in such processes, and can be used to obtain computational models that can explain many cognitive phenomena and situations. Combining knowledge of various neurocognitive findings and theories related to human awareness and control in behavioural choices, designing cognitive models based on such knowledge, and validating their behaviour are the main focus of this thesis. The scope of this work expands to various different but connected cognitive phenomena, namely: intentional inhibition, cognitive controlling, emotion generation, effect prediction, impact prediction, awareness generation, ownership generation, top-down effects, bottom-up effects, biased perception, cognitive conflicts, joint decision making, and analogy making. In addition to the contribution to dynamic computational cognitive models, this work includes two approaches for parameter estimation, which is essential in cognitive modelling.

The proposed models in this thesis have been applied to two practical domains: aviation and energy management. Simulation results that were generated for practical scenarios in those domains provided useful information about human cognition, while highlighting the strength and usability of the proposed models in practical context.
# Contents

## Part I: Introduction

<table>
<thead>
<tr>
<th>Chapter 1: Introduction</th>
<th>03</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Motivation</td>
<td>05</td>
</tr>
<tr>
<td>1.2 Research objectives &amp; research questions</td>
<td>08</td>
</tr>
<tr>
<td>1.3 Detailed overview of research questions</td>
<td>10</td>
</tr>
<tr>
<td>1.4 Methodology</td>
<td>25</td>
</tr>
<tr>
<td>1.5 Thesis outline</td>
<td>30</td>
</tr>
<tr>
<td>1.6 Personal contribution to each chapter</td>
<td>34</td>
</tr>
</tbody>
</table>

## Part II: Modelling the Role of Awareness, Emotion, Ownership and Control in Human Action Selection

<table>
<thead>
<tr>
<th>Chapter 2 - Computational Cognitive Modelling of Action Awareness: Prior and Retrospective</th>
<th>49</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction</td>
<td>50</td>
</tr>
<tr>
<td>2.2 Action Awareness Viewed Neurologically, Psychologically and Behaviourally</td>
<td>52</td>
</tr>
<tr>
<td>2.3 Description of the Cognitive Computational Model</td>
<td>57</td>
</tr>
<tr>
<td>2.4 Simulation Results</td>
<td>68</td>
</tr>
<tr>
<td>2.5 Discussion and Future Work</td>
<td>90</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 3 - Modelling Intentional Inhibition of Actions</th>
<th>103</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Introduction</td>
<td>104</td>
</tr>
<tr>
<td>3.2 Conceptual Basis</td>
<td>105</td>
</tr>
<tr>
<td>3.3 Description of the Computational Model</td>
<td>110</td>
</tr>
<tr>
<td>3.4 Analysis of the Model Based on Simulation Experiments</td>
<td>120</td>
</tr>
<tr>
<td>3.5 Discussion</td>
<td>129</td>
</tr>
<tr>
<td>Chapter 4 - Modelling the Dynamics of Emotional Awareness</td>
<td>139</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>140</td>
</tr>
<tr>
<td>4.2 Neurological Background</td>
<td>140</td>
</tr>
<tr>
<td>4.3 Description of the Model</td>
<td>142</td>
</tr>
<tr>
<td>4.4 Simulation Results</td>
<td>147</td>
</tr>
<tr>
<td>4.5 Discussion</td>
<td>151</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 5 - Modelling Dynamics of Cognitive Control in Action Formation with Intention, Attention, and Awareness</th>
<th>157</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Introduction</td>
<td>158</td>
</tr>
<tr>
<td>5.2 Cognitive, Behavioural, and Affective Science Evidence on Top-Down Guided Action Formation</td>
<td>159</td>
</tr>
<tr>
<td>5.3 Description of the Computational Model</td>
<td>162</td>
</tr>
<tr>
<td>5.4 Analysis of the Model Based on a Simulation Experiment</td>
<td>168</td>
</tr>
<tr>
<td>5.5 Discussion</td>
<td>171</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part III: Modelling Cognitive Metaphor in Joint Decision Making</th>
<th>175</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 6 - Modelling Analogy Making Based on a Simplified Buddhist Explanation</td>
<td>177</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>178</td>
</tr>
<tr>
<td>6.2 Motivation</td>
<td>178</td>
</tr>
<tr>
<td>6.3 Evolution of Analogy Making Models</td>
<td>179</td>
</tr>
<tr>
<td>6.4 Emergent Intelligence</td>
<td>185</td>
</tr>
<tr>
<td>6.5 Cognition through Five Aggregates</td>
<td>185</td>
</tr>
<tr>
<td>6.6 Approach and Design</td>
<td>187</td>
</tr>
<tr>
<td>6.7 Geometric Analogy Solver</td>
<td>189</td>
</tr>
<tr>
<td>6.8 Evaluation and Conclusion</td>
<td>191</td>
</tr>
</tbody>
</table>
### Chapter 7 - Modelling the Role of Cognitive Metaphors in Joint Decision Making

- **7.1** Introduction & Motivation
- **7.2** Core Concepts: Mirroring, Internal Simulation, Ownership, and Metaphors
- **7.3** A Computational Dynamic Social Agent Model for Joint Decision Making with Metaphors
- **7.4** Simulation of Example Scenarios
- **7.5** Discussion

### Part IV: Application of Integrative Computational Models in Complex Real-World Domains: Aviation and Domestic Energy Management

#### Chapter 8 - Parameter Estimation for Computational Cognitive Models

- **8.1** Introduction
- **8.2** Theoretical basis of CS and PSO
- **8.3** A Computational Cognitive Model of Action Awareness
- **8.4** Parameter Estimation with PSO
- **8.5** Discussion

#### Chapter 9 - Modelling Situation Awareness with Perception, Attention, and Prior and Retrospective Awareness

- **9.1** Introduction
- **9.2** Situation Awareness Related Processes Viewed Neurologically, Psychologically and Behaviourally
- **9.3** Description of the Computational Model
- **9.4** Specification of Model Compilation
- **9.5** Analysis of SA on the Proposed Model based on Simulations
- **9.6** Discussion
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.4</td>
<td>Integrated model validation through simulations</td>
<td>345</td>
</tr>
<tr>
<td>13.5</td>
<td>Discussion</td>
<td>350</td>
</tr>
</tbody>
</table>

**Part VI: Discussion**

**Chapter 14: Discussion**

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.1</td>
<td>Summary of research questions</td>
<td>358</td>
</tr>
<tr>
<td>14.2</td>
<td>Relations between models developed over time</td>
<td>366</td>
</tr>
<tr>
<td>14.3</td>
<td>The applicability of the research in this thesis</td>
<td>368</td>
</tr>
<tr>
<td>14.4</td>
<td>Future work</td>
<td>371</td>
</tr>
</tbody>
</table>

**සාරාංශය (In Sinhala)**

**SIKS dissertation series**

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>381</td>
</tr>
<tr>
<td></td>
<td></td>
<td>385</td>
</tr>
</tbody>
</table>
Part I:

Introduction

Although they may be complex, understanding of human cognitive processes is vital for many application domains (for example, health, aviation, energy, and safety critical domains) to develop innovative applications. Processes involved in human awareness and control have received much interest within the research community and various approaches have been proposed to understand and explain the working mechanisms of cognitive processes in the human brain. The fundamental research question of this thesis is how we can understand and explain human awareness and control in behavioural choices through dynamic computational models. For this purpose, relatively new knowledge from the neurocognitive research area (strongly fuelled by neuro-imaging and recording techniques) and dynamic system modelling methods have been used. Furthermore, the explored dynamic cognitive models have been validated in the context of socio-technical systems in aviation and energy domains. This thesis addresses the mechanisms involved in different cognitive elements relevant to the main theme, including awareness, emotion, ownership, perception, attention, intention, desire, control in human action selection, and metaphors. In addition, for the application domains aviation has been studied and energy management, in particular the performance of heat pumps in domestic heating related energy management are explored to integrate cognitive models into this application domain and to scrutinize the processes of such a hybrid socio-technical system.
Chapter 1

Introduction

The human brain is a complex, intricate, adaptive, dynamical system; it is difficult to unravel it and comprehend its mechanisms (Bressler, 1995; Fries, 2005; Friston, 2002; McIntosh, 2000; Mesulam, 1998; Swanson, 2012). Nevertheless, cognitive functionalities in the human brain and mechanisms underlying them are interesting to understand and those are great challenges to scientists, researchers and/or philosophers. The rapid developments in brain imaging and recording techniques (especially in the last three decades) allows the research community to uplift the understanding of brain processes and mechanisms while forming new branches of research to scrutinize these functionalities from different point of views (Posner & Raichle, 1999; Raichle, 2003; Kwong et al., 1992). Understanding the processes in the human brain demands a multidisciplinary approach. Cognitive modelling has increasingly become a prevalent multidisciplinary research theme for understanding those processes in the human brain and contributes to improving the knowledge of human cognition and facilitating more intelligent and realistic applications for many domains, among which Artificial Intelligence (AI) (Shiffrin, 2010; McClelland, 2009; Addyman & French, 2012).

Among many processes in the human brain, human awareness has received much attention from many researchers; e.g., (Baumeister, Masicampo, & Vohs, 2011; Desantis, Hughes, & Waszak, 2012; D'Ostilio & Garraux, 2012; Haggard, Clark, & Kalogeras, 2002; Haynes, 2011; Libet, Gleason, Wright, & Pearl, 1983; J. W. Moore, Lagnado, Deal, & Haggard, 2009; Moore & Obhi, 2012; Walsh & Haggard, 2013; Wegner, 2002). From a cognitive neuroscience and physiology perspective it is useful to understand causes behind some cognitive and behavioural disorders that relate to awareness. This leads to improvements in the current treatment options and, more generally, allows to improve mental health related aspects in more rational manner. From the AI perspective, it is interesting to understand the process behind natural intelligence and how we develop awareness and its contribution to decision making. Human awareness is a tacit term, but is recognised as an important phenomenon for human consciousness (Blackmore, 2010; (Bud) Craig, 2009; J. D. Cohen & Schooler, 1997). In particular, the interplay between conscious and unconscious processes is interesting to explore.

Among many roles of human awareness for cognitive functionalities, action selection has obtained a strong recognition within the research community (Baumeister, Masicampo, & Vohs, 2011; Haggard, Clark, & Kalogeras, 2002; Libet, Gleason, Wright, & Pearl, 1983; Walsh & Haggard, 2013). It is a common
belief among humans that we are aware of what we are doing and therefore we are responsible for our actions. However, it is a question how we actually select our actions and why we select that action in a given situation instead of something else (especially in complex situations). Is that purely a random selection? If not, how do we select a rational choice of action and how does awareness contribute to it? These are important questions to explore. The question of how and why human cognition emerges is not only about action selection, but also about action inhibition: sometimes we intentionally inhibit an action that already was decided. These issues highlight the importance of awareness for a social being. Furthermore, within a particular situation, action selection may differ from person to person significantly. Therefore, it is a challenging research question how human awareness contributes to action selection, what are the processes behind this, and to explain the differences in selected actions in a given situation.

Human awareness may have specific forms, for example, action awareness, emotional awareness, and situational awareness. Many applications (especially in safety critical domains and healthy lifestyle applications) lack knowledge about these different forms of human awareness. Identifying the mechanisms underlying human awareness can be quite helpful to design and develop more advanced and intelligent applications that will contribute to improve the quality of those applications. Often such applications need to have adequate understanding of the processes: they are required to be ‘human-aware’ or ‘socially aware’; e.g., (Pentland, 2005; Pantic, Pentland, Nijholt, and Huang, 2006; Treur, 2008; Bosse, Hoogendoorn, Klein, and Treur, 2009); see also, for example, relating to attention: (Bosse, Memon, Treur, and Umair, 2009; Bosse, Memon, Oorburg, Treur, Umair, and Vos, 2011). When the human processes considered in such applications also involve human awareness, then this opens the possibility to design applications that are human-aware, in particular for human awareness, or shortly, that are human-awareness-aware.

Most research in cognitive neuroscience focuses on relatively small but highly important questions using increasingly detailed data (Bassett & Gazzaniga, 2011). Therefore, it is difficult to find detailed explanations of a certain phenomenon (for example action awareness) as an integrated, coherent system. A process of one selected phenomenon always is affected by many others and vice versa, due to the high order of coupling and associations in the human brain. Therefore, there is room in brain related research areas to design and model compound systems by integrating various types of information collected from various neuro-cognitive research findings and theories. Having such a compound model can also be used as a workbench to scrutinize various hypotheses in cognitive neuroscience research (McClelland, 2009) and this can in turn contribute to improving the quality of empirical research, especially through the insights that can be collected through
interesting simulations. On the other hand, there are many implications of various hypotheses, theories or findings, and it would be beneficial if there was a mechanism that can be used to scrutinize these ideas by using this as a workbench at a much more abstract and global level (McClelland, 2009). Additionally, the human brain and its phenomena concern immeasurably complex systems and processes that involve uncountable many factors that make experiments not always coherent with reality. Nevertheless, having computational models enables to uplift the progress of understanding these processes in a broader level as a multidisciplinary approach (Shiffrin, 2010).

In this thesis, in addition to awareness many other cognitive states are considered, such as: perception, sensory representation, action preparation, predictive feeling, performatve desires, attention, subjective desires, intention, emotions, ownership, metaphors, action execution, and communication. A main focus is on dynamics of internal processes together with these interesting cognitive states, in contrast to social aspects of human cognition. Nevertheless, social aspects are also partly considered in this thesis through addressing cognitive metaphors in joint decision making. These models were scrutinized for aviation and energy domains to validate the model functionality for some selected real world situations/scenarios. In particular, to address the energy application domain, some additional research also involved models for energy related physical processes, especially concerning air to water heat pumps. This additional work provided good insight to combine energy models together with cognitive models to obtain more realistic simulation results.

1.1. Motivation

With the strong developments in brain-imaging and recording techniques, more and more detailed information on various brain processes becomes available. This contributes to an exponential increase in the development of (biologically inspired or brain-inspired) cognitive modelling (Bandettini, 2009). It has been found that nowadays more than 80% of the published articles in theoretical journals in cognitive science are about cognitive modelling (Busemeyer & Diederich, 2010). Furthermore, computational modelling is considered to be an important pillar for the development of cognitive science and its related disciplines (Addyman & French, 2012; McClelland, 2009; Shiffrin, 2010). In principle, in cognitive modelling, a phenomenon of human cognition (or behaviour) is represented as a (computational) mathematical model to understand causality, thereby using adjustable parameters that can be estimated on the basis of neuro-cognitive findings or theories (Addyman & French, 2012; Busemeyer & Diederich, 2010; Lewandowsky, 2011; McClelland, 2009; Shiffrin, 2010). These cognitive models provide more flexibility to deal with the complexity attached to human brain
Chapter 1

processes and contribute to the aggregation of many interesting findings from brain research. The relation between actual human cognition and empirical evidence collected through brain imaging and recording techniques bridges two different levels of abstraction. Therefore it is important to explore alternatives to handle this gap and to translate interesting and useful information into a different form that can be easily understandable and expressible. Dynamic modelling approaches provide insight for this and having a model that can explain many cognitive phenomena and situations will be useful in multidisciplinary context. For example this form of models can be used as a workbench at a much more abstract and global level in cognitive, behavioural, and affective sciences related experiments to scrutinize different hypotheses and to explore the mechanisms of different processes integrated as a coherent system.

There also exist well established cognitive architectures that are popular among specific groups of researchers; for example, (Chong, Tan, & Ng, 2007; Duch, Oentaryo, & Pasquier, 2008; Langley, Laird, & Rogers, 2009; Taatgen & Anderson, 2010). Some of such popular cognitive architectures are: ACT-R (Anderson & Lebiere, 1998), Soar (J. E. Laird, Newell, & Rosenbloom, 1987; J. Laird, Rosenbloom, & Newell, 1986), Clarion (Sun, 2005), LIDA (Franklin & Patterson Jr, 2006), and EPIC (Kieras & Meyer, 1997; Meyer & Kieras, 1997). These architectures are very useful and interesting, and also contribute to design and development of many complex applications with much cognitive detail in applications. Nevertheless, there is always a gap between sticking to a cognitive architecture (mainly based on a specific theory or hypothesis) and designing new dynamic models combining many different findings/hypotheses/theories, which are for example more recent and related to current neurocognitive evidence (Duch et al., 2008; Langley et al., 2009; Taatgen & Anderson, 2010). Therefore, there is a trend in designing dynamic cognitive models specifically combining the detailed micro-level information that is discovered through brain imaging related research. The main advantage of this type of information is that it includes a detailed level of information on very specific mechanisms in the brain. Therefore, it is possible to eliminate generic assumptions while providing strong abstract explanations on how the selected process works. In addition, due to the higher order coupling in brain processes, through these techniques it is possible to identify the relation between different cognitive processes and to validate their probable interplay. Also this further supports coherent, detailed effects of states on each other to facilitate modelling of their dynamics. Therefore, both well-known cognitive architectures and specific but more detailed dynamic cognitive models as presented in this thesis contribute to the development of multidisciplinary brain research.

Neurologically justifiable human-like models are useful for the development in many research domains including Cognitive (Neuro-)Science and (Human-Aware)
Artificial Intelligence. Nevertheless, achieving neurologically justifiable human-like models is a non trivial research challenge. There are well known models: Belief-Desires-Intention (BDI) (Georgeff et al. 1999), emotional BDI (eBDI) (Pereira et al. 2005; Jiang & Vidal 2006), Beliefs-Desires-Obligations-Intentions (BOID) (Broersen et al. 2002; Broersen et al. 2001); which are very popular and contributed to the development of AI (especially for multi agent system modelling). Nevertheless, a main question is in how far these are justifiable through neurological evidence. Certainly these have had much influence on the development of weak AI, but it may be doubted how such a type of models can contribute to the development of strong AI, as they are difficult to justify through neurological evidence. In general (simple) linear chain models or bus architecture based models have only limited features and are far away from real emergent brain processes in humans, and therefore may not reflect human-like behaviours. A main reason behind the complexity of processes in the human brain is the frequent occurrence of cyclic loops and coupling with other cyclic processes. Therefore, in the real world a cognitive process is mostly a coherent system based on a complex interplay of multiple interacting cyclic chains. Therefore, instead of linear models it is needed to take into account several aspects and (often cyclic) processes, which makes both the analysis of the required concepts and knowledge, and the models themselves more complex. There may be an impression or view that simpler models are always better, and more complex models are not really beneficial and their underlying knowledge which are used to develop those models also overly complex. However, complexity cannot be avoided when really human-like models are developed. Furthermore, when an appropriate modelling approach is used, complexity is also not a bottleneck but contributes to the study of the processes in a more broad manner and to understand the hidden effects and specially the cyclic nature of such processes. There are models in this thesis which in this sense are quite complex but still manageable and contribute to development of work-benches for neurologically justifiable human-like experiments and applications.

There are many domains where model-driven experiments are vital and unavoidable. The aviation domain is a good example for this and simulation driven situation analysis is a key line of research in this domain to analyse risks and safety related aspects. As it is very difficult to conduct actual experiments in safety critical domains such as aviation, model-based simulation methods provide necessary insight to achieve adequate knowledge about possible risks and safety related concerns in research experiments. Human cognition is an important factor in many aviation related scenarios and this holds especially for the topic of situation awareness (SA can be considered as a subjective quality or interpretation of the awareness of a situation a person is engaged in). The main limitation in such experiments is the lack of detailed cognitive models for processes related to human
awareness. Therefore, having sufficiently complex models, that provide detailed
cognitive information, are useful in such situations and these can be used instead of
human agents. Another good example domain is the energy domain. The various
phenomena in the energy usage domain include emergent features, due to the
dynamic nature of important factors (indoor temperature, outdoor temperature,
performance index of heat pumps, user behaviours, perception of level of comfort,
perception of energy cost, capacity of house, and etc.). Human involvement is also a
key factor in energy related scenarios and to use design experiments including
humans is always costly and time consuming (also it is difficult to expect the
necessary behaviour from them all the time). Therefore cognitive models are very
useful in some of the experiments of this type and this reduces the cost significantly
and contributes to provide more realistic data, depending on the performance and
quality of the cognitive model(s). Practical usage of these types of models is not
limited to these few examples, but these highlight the importance of such research.

1.2. Research objectives & research questions

The main research question of this thesis is:

How can dynamic computational cognitive models for human action
selection be designed, developed, simulated and applied, and what is the role
of awareness and cognitive control in such models?

This research question is broad and provides a more generic overview of the
scope of this research. Human cognition behind the process of action selection is the
main focus of this thesis, in particular in relation to awareness and cognitive
control. Furthermore, this research is not domain specific and can be used in any
domain that needs cognitive processes and behaviour specifically for action
selection. Nevertheless, two domains, namely aviation and energy, are selected in
this thesis to validate features of the designed models. Furthermore, a generic
challenge for nature-inspired dynamic models is parameter estimation. Especially
for cognitive models this is a non-trivial task, due to the difficulty of collecting
empirical data to align with (time series based) outputs generated from a model.
This issue was also considered within the scope of this research. Adhering to the
scope of the above research question, some aspects of it were highlighted in the
form of more specific research questions namely:

1) What is the role of awareness in human action selection?
   a) What is the role of unconscious and conscious processes and their
      interaction in human action selection?
   b) How does the internal prediction process shape or contribute to the
      (prior) awareness of the action?
c) How does inferential sense making shape or contribute to the (retrospective) awareness of action execution?
d) How does awareness contribute to action execution?

2) What are the roles of emotions, ownership and cognitive control in human action selection?
   a) What are the roles of action ownership and action awareness states in action selection?
   b) What can be learnt from cognitive/affective/behavioural sciences on how perception, attention, intention, emotion and awareness contribute to action formation and cognitive control?
   c) What can be learnt from cognitive/affective/behavioural sciences on how an emotion generation process interacts with action selection?
   d) How to model the interplay between conscious vs. unconscious processes in human action formation?

3) How to design cognitive models for joint decision making on action selection and what is the role of cognitive metaphors in such models?
   a) What can be learnt from Buddhist explanations for analogy making?
   b) What can be learnt from cognitive metaphors for joint decision making on action selection?

4) How can the developed computational models be applied in complex real-world domains?
   a) How to estimate parameters in dynamic cognitive models, especially given incomplete empirical features or patterns over time?
   b) How to embed nature-inspired human cognitive processes for situation awareness in safety critical domains?
   c) How to improve the analysis of domestic energy management for heating through dynamic computational models?
   d) How to utilize cognitive models in intelligent energy management through simulations to uplift the state of art in current system automation?

These questions are explored in this thesis, aligning with the scope of the main research question. The main research question focuses around human action selection together with the roles of awareness, and cognitive control. From an initial literature search it was observed that there is a debate about the contribution of awareness (consciousness-related) to action selection. This debate is the reason to have research question 1.a. Moreover, the contributions by Haggard and co-workers highlight the importance of distinguishing action awareness as prior and retrospective processes relative to action execution. This leads to the exploration of
the roles of internal prediction and inferential sense making for awareness of the action: research questions 1.b and 1.c. Starting from the beginning on how awareness directly contributes to action execution was an atomic question for this research. This question is included as 1.d.

In addition to action awareness, there are a number of other cognitive states that were found as important, mainly through literature search for question 1. Therefore, it was decided to explore emotions, ownership and cognitive control in human action selection related processes. One of the reasons to select these three elements was their strong relation to awareness. Emotions play a key role in many cognitive states and awareness states, and it is a useful concept in most human cognition related applications anyway. Furthermore, ownership and awareness also have strong relationships and prior models for the role of ownership in action selection also have been developed (Treur, 2012). The idea behind cognitive control was that it seems that there are complex situations which cannot be properly explained through the initial models developed for this thesis. For example how humans shift action choices in cognitive processing was a major question. Therefore the idea of cognitive control was considered in this work. These have led to four sub questions: 2.a, 2.b, 2.c, and 2.d.

The question on how to design cognitive models for joint decision making on action selection with a role for cognitive metaphor is an extension to the initial line of research to explore wider applicability. The other models in this thesis address the behaviour of a single human being. Nevertheless, many applications need the behaviour of groups, and joint decision making on action selection. Therefore, a separate research question was included to explore how to design cognitive models for joint decision making on action selection. When doing this, prior work that was done earlier was found relevant and also added to the scope. Analogy making based on Buddhist theories was research conducted earlier and addresses how different structures can be mapped onto each other and how this enables transfer of knowledge between two different contexts, just like how metaphors are mapped onto new situations. Finally research questions 4a, 4.b, 4.c, and 4.d were included to explore the applicability of the developed cognitive models in real world application domains. Furthermore, from the beginning of this research, it was clear that for applications, work on parameter estimation methods for dynamic computational cognitive models is needed, which is also included as a research question. This also explored under this applicability area.

1.3. Detailed overview of research questions

This thesis is composed of five parts where each part includes one or more chapters. Most chapters include published research and these contribute to answer one or more research questions as presented above. Therefore, each part reflects a
specific theme that contributes to the main research question. Part I is the current chapter of the thesis, which motivates the importance of dynamic computational cognitive modelling related research and describes the research questions together with an overall research methodology. Part II of the thesis addresses modelling the role of awareness, emotion, ownership and control in human action selection. This part is a collection of four chapters which include number of concepts and processes:

- Conscious and unconscious processes and their interaction in relation to action selection
- Prior and retrospective effects particular to action awareness and ownership
- Awareness generation and its effects on other processes to drive action selection
- Contribution of awareness on action execution
- Impact prediction processes
- Effect prediction processes
- Performative and constitutive desires in intentional inhibition
- Cognitive controlling
- Emotion generation

Part III covers modelling cognitive metaphors in joint decision making for action selection and a model for analogy making. Cognitive metaphors are important in joint decision making and having different metaphors leads to a cognitive conflict. This model was illustrated in particular for two types of metaphors that can affect joint decision making in different manners: a cooperative metaphor and a competitive metaphor. The way in which the role of metaphors was modelled was inspired by a model for analogy making adhering to a (simplified) Buddhist explanation. Part IV describes applications of integrative computational models in complex real world domains. As part of the process of designing and developing cognitive models it is important to evaluate and explain how to apply them in real world situations. Mainly two domains were selected for this purpose: aviation and energy. For the aviation domain situation awareness was simulated through a dynamic cognitive model that was developed. These simulations cover a number of different situations: poor perception (due to failure to correctly perceive information), incorrect comprehension (due to failure to rationally comprehend the situation), incorrect projection (due to failure to project a future situation properly), and a conflict between what is predicted and what actually occurs. For the energy domain, some research was conducted to evaluate the use of the type of models developed in heat pump related energy management. Therefore, an analytical model for mathematical analysis of smart daily energy management for air to water heat pumps was developed and it was separately validated with some real data too.
Having this analytical model (with separate validations) and a cognitive model for action selection, these two were integrated to provide insights for cognitive driven energy management choices. In addition, in all cognitive models an important challenge is parameter estimation. Therefore, an improved parameter estimation method was developed that is particularly useful for dynamic cognitive models for which sufficient empirical data cannot be collected. Part V provides an overall discussion about this research.

A summary of how each (sub) research question is reflected in the chapters of the thesis is as follows. The thesis consists of 14 chapters and except the Introduction and Discussion almost all chapters are based on published conference or journal papers (all are published work except Chapter 6).

<table>
<thead>
<tr>
<th>Part</th>
<th>Chapter</th>
<th>Research Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>2</td>
<td>1.a, 1.b, 1.c, 1.d, 2.a</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.a, 1.b, 1.d</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.d, 2.a, 2.c, 2.d</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.a, 2.b, 2.d</td>
</tr>
<tr>
<td>III</td>
<td>6</td>
<td>3.a</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>3.b</td>
</tr>
<tr>
<td>IV</td>
<td>8</td>
<td>4.a</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.c, 1.d, 2.a, 2.b, 4.b</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>4.c</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>4.c</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>4.c</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>4.d</td>
</tr>
</tbody>
</table>

### 1.3.1. Research Question 1.a What is the role of unconscious and conscious processes and their interaction in human action selection?

Human action selection is a complex process and the role of conscious and unconscious processes in this is an important issue. Chapter 2 explores a computational cognitive model for action awareness. Some evidence leads to a hypothesis that awareness of action selection is not directly causing the action execution (or behaviour) but comes afterward, as an effect of unconscious processes of action preparation (Baumeister, Masicampo, & Vohs, 2011; D’Ostilio & Garraux, 2012; Haynes, 2011; Libet, Gleason, Wright, & Pearl, 1983; Wegner, 2002). In contrast, another hypothesis claims that both predictive and inferential processes related to the action preparation and execution may contribute to the conscious awareness of the action and furthermore, this awareness of an action is a dynamic combination of both prior awareness (through predictive motor control processes) and retrospective awareness (through inferential sense-making processes) relative to the action execution (Desantis, Hughes, & Waszak, 2012; Haggard, Clark, & Kalogeras, 2002; Moore, Lagnado, Deal, & Haggard, 2009;
Chapter 1

Moore & Obhi, 2012; Walsh & Haggard, 2013). The presented model integrates findings on both conscious and unconscious explanations for both action awareness and ownership and acts as a generic computational cognitive model to explain behaviour through the interplay between conscious and unconscious processes.

Chapter 3 focuses on the process of intentional inhibition, i.e., the capacity to voluntarily suspend or suppress an action. This provides good insight in what is the role of unconscious processes in action selection and how difficult it is to suppress an action which is to be executed without the support of additional conscious processes (Brass & Haggard, 2007, 2008; Filevich, Kühn, & Haggard, 2012; Haggard, 2008; Kühn, Haggard, & Brass, 2009; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012). In an intentional inhibition process an internally guided action is inhibited using an internally guided stop signal. The interplay between (positive) potential selection of an action and (negative) predicted impact of this action is considered for this. According to neurological evidence it is possible to trigger a (positive) potential selection of an action mainly from the unconscious process, but it is essential to have conscious elements to enable generation of a (negative) predicted impact that leads to an action inhibition. This interplay of positive and negative influences on action selection has contributed to make homo sapiens into a social being. In this process performative and constitutive desires are used to differentiate the influence for action preparation and intentional inhibition, respectively, with the relevant supportive states: ownership and awareness states.

Chapter 5 further contributes to explain the role of unconscious and conscious processes in human cognition for action selection. This chapter focuses on cognitive control in action formation with intention, attention, and awareness. The presented model shows the interplay between the bottom-up and the top-down processes and how they interact with each other to cognitively drive a given situation to a probable solution, thereby taking into account attention, intention, and awareness (Engel, Fries, & Singer, 2001; Haggard, 2008; Kiefer, 2007; Miller & Cohen, 2001; Moore & Haggard, 2008; Rigoni, Brass, Roger, Vidal, & Sartori, 2013). The bottom-up processes are mainly due to unconscious effects, whereas top-down processes contribute to conscious influence. The model shows how the interplay of these two is the basis for generating the most appropriate response. Furthermore, additional inhibition processes of both conscious and unconscious types are also explored in this model, aligning with action selection specific mechanisms.

1.3.2. Research Question 1.b How does the internal prediction process shape or contribute to the (prior) awareness of the action?

Awareness of an action is a dynamic combination of both prior awareness (through predictive motor control processes) and retrospective awareness (through
inferential sense-making processes) relative to the action execution (Brass & Haggard, 2007, 2008; Filevich et al., 2012; Haggard, 2008; Kühn et al., 2009; Walsh et al., 2010; Zhang et al., 2012). In particular predictive processes (through internal simulation; Damasio, 1999, 2005, 2012) related to action preparation may contribute to the conscious awareness of the action, via the emotions they generate. This form of awareness is referred to as prior awareness (awareness that emerges before action execution). It is affected by other states that together lead to the emergence of prior awareness; in turn its effect on other states is also discussed in neurocognitive research. In this regard internal prediction processes (internal simulation of the action) play a key role in shaping or at least contributing to prior awareness. Research by Haggard and co-workers provide much evidence for how the internal prediction process shapes or contributes to the (prior) awareness of the action: (Brass & Haggard, 2007, 2008; Haggard, 2008; Haggard, Clark, & Kalogeras, 2002; Walsh & Haggard, 2013). One specific occurrence of this prediction process is based on Damasio’s as if body loops (Damasio, 1999, 2005, 2012), which predict particular body states. Chapter 2 explains in more detail how predictive processes of action execution play a role in conscious awareness of an action. The cognitive process of action selection is based on valuing the outcome of an internal simulation process prior to the execution of an action (Damasio, 1999, 2005, 2012). The brain will evaluate the effect of each relevant action option by comparing the feelings associated to each individual predicted effect (without actually executing them). The simulated option that has the strongest valuated feeling performs as a GO signal for execution of that action and the other options are NO-GO options. This process plays an important role in the emergence of action ownership (see Treur, 2012). Lateral inhibition processes (Aron, 2007) contribute to further strengthening of the action selection process. Therefore, naturally the strongest internally satisfactory option (which is exceeding a threshold value) will become selected as a result of the unconscious action selection process. This strong prior ownership and the predicted feeling seem to contribute to the emergence of prior action awareness as well.

This internal prediction processes is further explored in Chapter 3 by specifically considering two different predictive processes: effect prediction and impact prediction. In this contribution impact prediction is a new addition in comparison to Chapter 2. This impact prediction is triggered as an effect of the unconscious effect prediction process. The effect prediction processes concerns more the positive habitual aspects of an action selection. Nevertheless, it is not always possible to let habitual aspects exclusively drive the cognitive action selection processes, especially when surviving as a ‘social entity’ is considered (J. R. Cohen, Berkman, & Lieberman, 2013; Mostofsky & Simmonds, 2008). Impact prediction includes so-called constitutive desires (which drive a person to his or her
long term driven aspirations) to inject more conscious aspects to the processes and provides more subjective influences together with prior awareness. Therefore it evaluates the possible negative influence of the current action selection from the longer term perspective and that may lead to abandoning the action with prior awareness. Performative desires (which represent short term interests/goals) and constitutive desires play a key role to provide the above mentioned positive and negative evaluations.

1.3.3. **Research Question 1.c** *How does inferential sense making shape or contribute to the (retrospective) awareness of action execution?*

Similar to the internal prediction process, inferential sense making processes are also important for the awareness of action execution. This contributes to the differentiation of cognitive effects before and after action execution. This difference is very important, especially when there is a mismatch between what was predicted and what really occurred. An awareness state can also develop in retrospect, after the action was performed. Haggard and co-workers point out that awareness of an action is a dynamic combination of both prior awareness (through predictive motor control processes) and retrospective awareness (through inferential sense-making processes) relative to the action execution (Desantis, Hughes, & Waszak, 2012; Haggard, Clark, & Kalogeras, 2002; Moore, Lagnado, Deal, & Haggard, 2009; Moore & Obhi, 2012; Walsh & Haggard, 2013). Retrospective awareness answers the question: ‘*what have I done?*’. Such a retrospective awareness state often relates to acknowledging others and taking responsibility for having performed the action. It may also play an important role in learning based on experience: by evaluating the obtained effect in a conscious manner leads to improvement of the performance of the action selection in the future. These retrospective processes are considered in Chapter 2 where it is explained how they shape the awareness of action execution. Chapter 9 includes a more improved dynamic cognitive model and specifically considers simulating a situation that shows a conflict between what is predicted and occurred. In this model, when there is a mismatch between what is predicted and what is actual, this mismatch can be considered as represented through prior and retrospective awareness states. More specifically, when, while predicting a strong satisfactory effect and the agent develops strong prior awareness, the same predicted effect is not sensed afterwards, the agent will develop poor retrospective awareness of that predicted effect. Through this mechanism, the agent can examine the performance of an executed action relative to its predicted effects in a particular situation. Having such a mechanism to interpret a mismatch between predicted and occurred effect is important in many situation awareness scenarios. More importantly, through this process, the agent can learn to consider possible
alternative options by having realised that an action does not lead to a particular effect as expected.

1.3.4. **Research Question 1.d** How does awareness contribute to action execution?

When a prior awareness state occurs, a person may have become aware of going to perform the action. Having such a prior awareness state may leave open the question whether the agent is able to consciously decide to perform or not to perform the action (Baumeister, Masicampo, & Vohs, 2011; Custers & Aarts, 2010; D’Ostilio & Garraux, 2012; Haynes, 2011; Libet, Gleason, Wright, & Pearl, 1983; Wegner, 2002). Moreover, it is equally doubtful whether all actions can be executed or suppressed without such awareness. For example, is still some form of vetoing of the action possible, without having prior awareness? In principle, the prior awareness state may play the role of generating a kind of green light for execution of the action. However, equally well the prior awareness state may only have an internal monitoring function and just play the role of a warning for the agent to be prepared that the action will happen (anyway). As stated in (Baumeister, Masicampo, & Vohs, 2011; Haynes, 2011; Libet, Gleason, Wright, & Pearl, 1983; Wegner, 2002) it has been found that for certain types of actions the decision to perform it is already made at least hundreds of milliseconds (and even up to 10 seconds) before any awareness state occurs. These findings may suggest that prior awareness often will have no effect on the decision. But this may strongly depend on the type of action. For example, it will be difficult to believe that the action of buying a car or a house remains unconscious and may not be amendable to vetoing based on awareness states (the Monty Hall Problem is also another good situation for this concern). Therefore, models for cognitive processes were designed considering both claims: awareness has an effect on action execution and it has not.

Chapter 2 provides a perspective on the basic processes adhering to the claims presented above. Through bottom-up activations the agent develops prior awareness and through this the agent can inject some bias to the current unconscious processes. This may strengthen a weaker action option and improve the predictive feeling of that option, which may lead to getting it executed. Furthermore, in this model the prior awareness can also directly strengthen the action execution state. In Chapter 3 the model designed in Chapter 2 is further extended and illustrates how awareness is important in intentional inhibition. This contributes many factors to this research question and shows how difficult it is to inhibit an action intentionally when there is no awareness. Emotions also play a key role in action execution, and emotional awareness is also special form of awareness. Chapter 4 specifically focuses on how emotional awareness affects action awareness. This chapter incorporates a role for emotional awareness states with attention, and perception
that act reciprocally and interactively in the dynamics (top-down) of emotion generation, but also cover automatic, unconscious emotion generation processes (bottom-up), and the mutual interaction between these bottom-up and top-down processes (McRae, Misra, Prasad, Pereira, & Gross, 2012; Ochsner et al., 2009; Weinberg, Ferri, & Hajcak, 2013). Bottom-up emotion generation processes are elicited largely by emotional perceptions with weaker subjective aspects and not necessarily being conscious, whereas top-down processes go more together with conscious and appraisal processes driven by attention. Therefore emotion awareness through top-down processes is able to inject some bias to the action preparation process that directly affects action execution. Chapter 9 provides more information, especially for situation awareness. Poor situation awareness scenarios (from the point of perception, comprehension, projection, and mismatch between the predicted and actual effects of an action) are considered for this. These simulations provide detailed information about when and how awareness contributes to action execution. Except the wrong perception, other situations directly coupled with awareness and the affect and effects of this are discussed in this chapter.

1.3.5. Research Question 2.a What are the roles of action ownership and awareness states in action selection?

Action ownership and action awareness are considered two distinct cognitive states which have some relationship. Action ownership is a useful concept which is mainly important to differentiate in how far a person attributes an action to him or herself, or to another person (Treur, 2012). Although in many cases the feeling that you get when you perform some action or when another person is performing the same action may be similar, it is clearly possible to identify whether the action belongs to you or to someone else. More importantly, the information about another person’s behaviour influences your self-evaluation and vice versa, which makes humans social beings (Chaminade, Marchant, Kilner, & Frith, 2012; Decety & Sommerville, 2003). This separation of self and other is contributing to the ability to recognise ownership of your own action. Research has shown evidence that action prediction (based on sensory information) plays a crucial role in action execution with ownership; when there are problems with action prediction that often leads to abnormal states of ownership of that action. Furthermore, this ownership further contributes to the development of awareness together with the predictive and inferential feeling. This process is explained in Chapter 2 and is illustrated by simulation results.

The roles of ownership and awareness together with emotions in action selection are further studied in Chapter 4. The model discussed in Chapter 2 is further strengthened by extending it with more cognitive states to get a more detailed picture of the causality within the action selection process. The interplay between
bottom-up and top-down processes are the main extension of this model and through these new processes the effects on ownership and awareness are explained in a more detailed manner, and illustrated by simulation-based validations. This leads to a discussion on how far ownership and awareness can change (or dilute) emotion-inspired action and how hard this is when the emotions are very strong, with a strong perception. Chapter 9 also contributes to this research question and addresses how cognitive control related processes have effect on ownership and awareness have. Evidence from literature and the designed cognitive models provide a better explanation for the role of ownership and awareness on action selection. Specifically the clear separation of prior ownership and awareness helps to differentiate effects on action execution.

1.3.6. Research Question 2.b What can be learnt from the (neuro-) cognitive / affective / behavioural sciences on how perception, attention, intention, emotion and awareness contribute to action formation and cognitive control?

Due to the rapid development in brain imaging and recording techniques there are many research findings available particularly from the (neuro-)cognitive, affective, and behavioural sciences. These findings indicate the importance for action formation of some cognitive states in particular: perception, attention, intention, emotion and awareness. Chapter 5 addresses those cognitive states and processes related with these findings from research of (neuro-) cognitive, affective, and behavioural sciences. The brain’s circuits for cognitive control seem to consist of loops rather than linear chains (Haggard, 2008). The prefrontal cortex (PFC) plays an important role in top-down driven cognitive control, as a temporal integrator (Miller & Cohen, 2001). The higher order interconnectivity of the PFC with other cortical, and subcortical areas has been interpreted as indicating a process that generates and maintains information when sensory inputs are weak, ambiguous, rapidly changing, novel and/or multiple options exist (Miller, 2000; Miller & Cohen, 2001). This chapter considers literature about cognitive control of action formation with circuits of perception, attention, intention, emotion and awareness. Chapter 9 is an extension of Chapter 5 but includes more detailed information on biased perception through perceptual load. Why and how people get distracted from their current task is explored here, specific to perceptual processes. Once a person is focusing on something, there are many reasons why he/she may sometimes be distracted and sometimes not. To explain this phenomenon, the load theory of attention and cognitive control (Lavie & Tsal, 1994) provides detailed information about early versus late selection schemes. An agent under high perceptual load is unable to shift his or her selection to other salient features in the environment.
Instead, when the perceptual load is lower, the agent is capable of perceiving information in parallel (Lavie, 2005, 2006; Lavie, Hirst, de Fockert, & Viding, 2004). This perception has direct effects on action formation and neuro-cognitive evidence behind this discussed in this chapter.

1.3.7. Research Question 2.c What can be learnt from cognitive/affective/behavioural sciences on how an emotion generation process interacts with action selection?

Emotions are vital and influential phenomena in human action selection. Therefore, it is important to explore how emotion generation process interacts with the action selection process from the perspective of (neuro-) cognitive/affective/behavioural sciences. Emergence of emotions has different explanations, varying from automatic responses (bottom-up), to more consciously emerging processes (top-down). These explanatory approaches have been able to explain emotional formation in line with results from fMRI experiments (Gross, 2002; Ochsner et al., 2009). Bottom-up emotion generation is assumed to occur immediately and ingrained from an external stimulus while top-down emotion generation involves semantic evaluation of a situation through cognitive influences (Sheppes & Gross, 2011). It has been shown that different neural areas are activated for this: thalamus, hypothalamus, ventral striatum, amygdala, anterior cingulate cortex (ACC), anterior insular cortex (AIC), orbito-frontal cortex (OFC), and/or mesial prefrontal cortex (Lane et al., 1998), with and without conscious intervention. The amygdala is the main hub not only for monitoring the emotionally salient stimuli but also for projecting to the relevant brain areas and transmit retrieved feedbacks to the sensory pathways, to invoke rapid and efficient generation of emotions (Brosch, Scherer, Grandjean, & Sander, 2013; Pessoa, 2010; Pourtois, Schettino, & Vuilleumier, 2013). According to (Pessoa, 2010) the amygdala directly shapes the perception when perceiving an emotionally salient stimulus (bottom-up) and by (Phillips, 2003) it is shown that emotional perception contributes to identify emotionally salient information in the environment, and to generate emotional experiences and behaviour. Also according to (Pourtois, G., Schettino, A., & Vuilleumier, P. (2013)) emotions can be shaped by perception through amplification mechanisms that do not overlap with other attentional processes (without leading to awareness). In this bottom-up process the brain has shown to capture the emotional perceptual features of the stimulus spontaneously without involving conscious awareness and further subjective aspects of this emotion (McRae, Misra, Prasad, Pereira, & Gross, 2012). Therefore, perception may drive emotion generation in the bottom-up approach.

Attention is a key cognitive process that allows (by subjectively desiring) to appraise a situation with conscious awareness (Suri, Sheppes, & Gross, 2013). While perception is a key aspect in the bottom-up process, attention compels the
20  Chapter 1

top-down process. Furthermore, there are mainly two types of attention mechanisms: exogenous (for bottom-up) and endogenous (for top-down); with partly distinct brain circuits (Pourtois, G., Schettino, A., & Vuilleumier, P. (2013); Weinberg, Ferri, & Hajcak, 2013).

Through these complex processes mainly involving attention and perception it seems like emotion generation directly interacts with action selection and provides necessary directions to decide which action to be selected. Chapter 4 covers the body of knowledge for this research question.

1.3.8. Research Question 2.d How to model the interplay between conscious vs. unconscious processes in human action formation?

The process of human action formation is a complex system that includes the influences of both conscious and unconscious processes. Having much neurocognitive evidence for this, it is important to exploit this body of knowledge by designing one compound model that explains this more abstractly but is in line with empirical evidences/hypothesis. Chapters 4 and 5 provide specific cognitive models (both are sharing the same basis but Chapter 4 focuses on emotional awareness whereas Chapter 5 focuses on cognitive control related to action formation) which includes the interplay of these two concepts. In this model action formation emerges more from unconscious processes and these include bottom-up processes (Katsuki & Constantinidis, 2014; McRae, Misra, Prasad, Pereira, & Gross, 2012; Ochsner et al., 2009; Pessoa, 2010; Pourtois, G., Schettino, A., & Vuilleumier, P. (2013); Sheppes & Gross, 2011; Weinberg, Ferri, & Hajcak, 2013). Having bottom-up processes this will pass information to high order cognitive states and enables top-down processes (Baluch & Itti, 2011; Engel, Fries, & Singer, 2001; Haggard, 2008; Katsuki & Constantinidis, 2014; Kiefer, 2007; Miller, 2000; Miller & Cohen, 2001; Ochsner et al., 2009; Pourtois, G., Schettino, A., & Vuilleumier, P. (2013); Sheppes & Gross, 2011; Tallon-Baudry, 2012; Weinberg, Ferri, & Hajcak, 2013). The model shows how the interplay of these two as a basis for generating the most appropriate response and how they interact with each other to cognitively drive a given situation to a probable solution using attention, intention, and awareness. By including the aspects of performative and subjective desires, perception, emotion, feeling, ownership, attention, intention, and awareness, this model provides an adequate level of detail (or complexity) to be used as a coherent system for various experimental needs to analyse behaviours. It may be beneficial as a work bench for some hypotheses in (neuro-) cognitive, affective, and behaviour sciences.
1.3.9. **Research Question 3.a What can be learnt from Buddhist explanations for analogy making?**

Analogy making is a fundamental process of the human cognitive process that helps to understand new things based on what already has been understood. Specially for creative thoughts this has been realised as very important by many scientists and understanding and modelling analogical thought process is one of the most important challenges for cognitive science (Mitchell, 1993). This research exploits the Theravada Buddhist theory of Five Aggregates to model the emergent intelligence through analogy making by meeting of related conditions (Thera, 1972, 2008). Investigating the role of five aggregates in the cognitive process to explain the analogy making is the main research question of this work. The Five Aggregates comprise of: Form, Sensation, Perception, Mental-formation and Consciousness (Boisvert, 1995; Thera, 2008). Form is the input to the mind while sensation will facilitate its relevant acceptance of it. Perception is about identification of the input and Mental Formation constructs the hypothesis together with the use of conditions that stem from already gathered knowledge. Consciousness is nothing but the ready-state of the mind. Chapter 6 provides details of what can be learnt from Buddhist explanations for analogy making and a cognitive model that was developed to explain the process of analogy making. This model was validated in the geometry domain.

1.3.10. **Research Question 3.b What can be learnt from cognitive metaphors for joint decision making?**

Making decisions together with others is an essential part of human life, especially in social and professional context. The way we understand many phenomena in our daily lives is metaphorical: one certain mental domain is understood in terms of another mental phenomenon. It is interesting to explore what can be learnt from cognitive metaphors for joint decision making in particular. In Chapter 7, the influence of cognitive metaphors on joint decision making has been examined by combining the concept of a joint decision making process and cognitive metaphor together in a computational social agent model. In this work two core processes were identified as important to model processes of joint decision making: mirror neurons (Iacoboni, 2008a; Treur, 2011a, 2011b) and internal simulation (Wolpert, 1997). Mirror neurons are motor neurons that fire when an action is (to be) executed by a subject, but also when the subject observes somebody else performing that action. The discovery of mirror neurons originates from single cell recording experiments with monkeys (Iacoboni, 2008a). After that discovery it has been subject to debate whether they also exist in humans. Due to limitations in temporal and/or spatial resolution of the brain imaging techniques used it was difficult to validate this through single cell experiments (Hickok, 2009).
In recent years the existence of mirror neurons in humans has still found support in single cell experiments with epilepsy patients undergoing pre-surgical evaluation of the foci of epilepsy (Fried, Mukamel, & Kreiman, 2011, Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010, Iacoboni, 2008a, Iacoboni, 2008b, Keysers, & Gazzola, 2010, Kilner, Neal, Weiskopf, Friston, & Frith, 2009). Therefore, this empirical evidence supports the existence of mirror neurons in the human brain and therefore can be used in cognitive models to strengthen the action preparation process. Internal simulation works in combination with mirror neurons. While mirror neurons prepare for an action, internal simulation generates a prediction of the (expected) effects of such a prepared action. Having these two concepts (i.e., mirror neurons and internal simulation) as an integrated process leads to empathic understanding (Damasio, 2005; Goldman, 2006; Hesslow, 2002) as part of joint decision making, thereby integrating the effect of ownership (Moore & Haggard, 2008; Treur, 2012). Having all these processes as underlying processes in joint decision making, it has been found that our metaphorical image of the situation have a strongly influence on this. The model integrates the mechanisms of cognitive metaphors into this process (Cardillo, Watson, Schmidt, Kranjec, & Chatterjee, 2012; Carroll & Thomas, 1982; Leary, 1994; Romero & Soria, 2005); these mechanisms were partly inspired by the way in which analogy making was addressed. Two cognitive metaphors (namely: a cooperative and competitive metaphor), are explored in a detailed manner in this chapter.

1.3.11. Research Question 4.a How to estimate parameters in dynamic cognitive models, especially given incomplete empirical features or patterns over time?

A complex phenomenon can be modelled in terms of the dynamics of its states and interactions, with temporal causal relations (cause and effect relations) as a basis for these interactions. A collection of such assumed causal relations can be compiled in the form of a dynamical model, which is more open to analysis than the real world phenomenon. A model is a close approximation, but always a form of abstraction of a real world phenomenon. Its accuracy and correctness mainly depend on the chosen abstraction assumptions and values of parameters in the model. Estimation of parameter values for a given model is a nontrivial task and it is much harder when it is a dynamical model. This is due to the fact that most of the cognitive models are dynamical models and only a limited amount of empirical data to validate those behaviours is available to conduct a specific parameter estimation process. Furthermore, it is difficult to adapt common parameter estimation algorithms to this situation that for cognitive models usually only limited empirical data are available (i.e., due to limitations in neuro-imaging techniques and the complexity of the human brain). Chapter 8 presents a solution for such cognitive
models using an improved Particle Swarm Optimization (PSO) (Kennedy & Eberhart, 1995; Palupi Rini, Mariyam Shamsuddin, & Sophiyati Yuhani, 2011; Poli, Kennedy, & Blackwell, 2007) algorithm combined with Constraint Satisfaction (CS) (Kumar, 1992; Tsang, 1996). In this approach, a parameter estimation problem for a given dynamic computational model is represented as a constraint satisfaction problem and an improved PSO algorithm is used to search for a possible solution together with identified partial temporal and static features and patterns.

1.3.12. **Research Question 4.b How to embed nature-inspired human cognitive processes for situation awareness in safety critical domains?**

Having a generic cognitive model for a specific type of process is a useful starting point to explain a real world phenomenon involving that type of process, which without such a model would be a nontrivial challenge. Specifically, having pre-developed models for action formation, cognitive control, and awareness states in action selection based on fundamental knowledge about processes in the brain, provides the basic ingredients to explain a given practical process in situation awareness from the perspective of action selection. Situation awareness is considered as a subjective quality or interpretation of the awareness of a situation a person is engaged in (Endsley, 1988). When a person is engaged in a situation based on the information that he/she perceives, the attention that is allocated to that information based on his/her subjective desires will develop his/her subjective awareness of the situation. Chapter 9 provides a detailed explanation for situation awareness through a neurologically inspired cognitive model. This research considers the process that Endsley suggested informally for SA. Nevertheless, according to the latest neuro-cognitive research there is a difference between the process that Endsley suggested and the latest findings. Therefore, this model includes the relatively new findings related to SA while giving attention to the key three levels that Endsley proposed in SA: perception, comprehension, and projection. This research includes four main scenarios in the context of a socio-technical system in aviation: a failure to correctly perceive information, a failure to comprehend the situation, a failure to project a future situation properly, and a conflict between what is predicted and what actually occurs. Having interesting results for simulations for 4 practical scenarios this provides good insight how to embed nature-inspired human cognitive processes for situation awareness in socio-technical systems in critical technical domains. For this chapter the aviation domain was selected, though the model is domain independent.
1.3.13. Research Question 4.c How to improve the analysis of domestic energy management for heating through dynamic computational models?

Domestic energy management is another challenging and interesting technically oriented research area (European Commission, Joint Research Centre, & European Technology Platform on Renewable Heating and Cooling (RHC-Platform), 2011). How to improve the analysis of domestic energy management for heating through dynamic computational models is a well-known question in the field. Moreover, this question also applies to the socio-technical system for domestic heating in which humans and technical equipment have to cooperate. Due to the cost and complexity in conducting practical experiments it is beneficial to design dynamic computational models that cover the behaviour of dynamic changes of environmental factors together with the behaviour of relevant energy generating systems (e.g., a heat pump). For this purpose behaviour of air to water heat pumps (Aste, Adhikari, & Manfren, 2013) together with the behaviour of indoor and outdoor temperature was modelled as a mathematical dynamical system. Chapter 10 presents an analytical model for mathematical analysis of smart daily energy management for an air to water heat pump. In this analytical model three phenomena are integrated: daytime and night-time outdoor air temperature, indoor temperature when cooling down takes place, and performance of a heat pump. The proposed approach calculates the energy used to heat a house according a particular heating program. The analytical model was compared with a simulation model to validate its accuracy. Chapter 11 is a continuation of Chapter 10 where an analysis of annual energy usages for domestic heating based on a heat pump was addressed. This uses the mathematical equation to calculate energy usage of a given date (considering the dynamic factors that influence to this: characteristics of the environment, characteristics of the house, and characteristics of the heating system) of an air to water heat pump. Having noted a simple but sensitive mathematical model this enables to simulate the energy usage predictions that can be compared with empirical data to validate the model parameters. Through this it is possible to get a good overall idea about probable energy usage of a year combining the trends of outdoor air temperature variations. Having a model to predict the energy usage more from analytical perspective Chapter 12 focus on how to estimate characteristics of a house based on sensor data. Domestic energy management should be adapted to the specific characteristics of each house. Having an analytical model for mathematical analysis this facilitates to integrate empirical data such that through the reverse engineering to isolate the characteristics of a given house.
1.3.14. **Research Question 4.d** How to utilize cognitive models in intelligent energy management through simulations to uplift the state of art in current system automation?

Human behaviour is not always easy to predict and may be complex. This is even more so if the environment in which the human functions is complex and dynamic. One example of such a complex and dynamic environment is the socio-technical system of domestic heating with the dynamics associated to indoor and outdoor air temperatures, required comfortable temperature set points over time, parameters of the heating source and system, and energy loss rate and capacity of a house. It is difficult to conduct real world experiments to analyse the dynamics and optimal efficiency of a heating system in actual daily use under all circumstances and constraints.

It is not only that cognitive modelling is a multi-disciplinary research but also its applications involve multiple application domains. Therefore, having Chapter 9 that explains how to apply a dynamic computational cognitive model into aviation domain (which more goes with cognitive science research) it is interesting to check whether such a cognitive model can also be applied to another application domain (e.g., the domestic energy management domain) that is not strongly related to cognitive science related disciplines. Nevertheless, in any research domain if human intervention is involved (in the core process) then it includes human cognitive factors. Therefore having 3 chapters (10, 11, and 12) that specifically address domestic energy management related analysis, Chapter 13 presents a research for cognitive simulation driven domestic heating energy management system using an air to water heat pump. It has integrated the earlier developed mathematical model for this heating system’s performance, and a computational dynamic cognitive model for the human’s behaviour which was developed based on evidence from Cognitive Neuroscience. Through this it addresses the importance of the choices of human behaviour on energy usage with a realistic setup through an alternative approach: a simulation-based analysis. Through this it gives an overall idea on possible impacts and provides additional information on how to motivate persons or correct their biased perceptions in a methodological approach to support lifestyle and lifestyle change in relation to energy management.

### 1.4. Methodology

Complex systems related research is often nontrivial and it is essential to have a well-defined methodology to conduct this. In this research a dynamic modeling approach is used which is influenced by (Ashby, 1978; Port & Van Gelder, 1998). The main perspective of this methodology is to model cognitive processes through networks of temporal-causal relations (Treur, 2016). Due to the rapid development
in brain imaging and recording techniques (Kwong et al., 1992; Posner & Raichle, 1999; Raichle, 2003) there is much research evidence available particularly for (neuro-) cognitive, affective, and behavioural sciences that could be used as fundamental basis for conceptual models ‘from first principles’. This knowledge was represented by temporal-causal networks that provide unambiguous but dynamical flow of a process. Nevertheless, this approach is different from neural network driven modelling and our focus was to realize the relation of key states with neurocognitive evidences more than mapping very detailed brain areas as states in the model. This ‘networks of temporal-causal relations perspective’ was chosen as a point of departure and it has been found that through this perspective it is relatively easy to design a model mainly at a conceptual, graphical level and to relate a model to scientific literature (Treur, 2016). In this design approach mainly two elements are used: nodes and arcs. Nodes are to specify cognitive entities (for e.g., perception, feeling, ownership, awareness, attention, intention, etc.) and nodes have inputs and outputs through arcs. More specifically, arcs allow linking with other states and external inputs and outputs (for, e.g., stimuli and actions). This is referred to as the conceptualisation phase of this methodology. The main challenge in this model design phase is to integrate different research findings and hypotheses into a compound model. Having many cyclic loops and higher order coupling makes this task difficult but it is important to maintain the generic nature.

Having a realistic cognitive model, it is essential to translate that into a computational form. This is referred to as the process of formalisation. For this purpose first a designed model is translated from graphical representation to a matrix representation. A description of the conceptual representation of a dynamical model in the first place involves representing states and connections between them, representing causal relations between states. Furthermore, this should include three main things to explain this as a dynamic computational process. The first is the way of specifying the strength of a causal connection. The dynamics can be effects through the strengths of each connection and it is essential to include this characteristic as a parameter. Each connection has a weight value representing the strength of the causal relation, often between 0 and 1, but sometimes also below 0 (negative effect). Having different values for connections this may lead to different behaviours by which certain characteristics of humans and situations/scenarios can be represented. The second factor is the method to combine causal impacts on a state. A state can be affected from one or many inputs and therefore it is essential to have a method to integrate all the inputs coming to a state to generate a unique output (which is sensitive to the strength of each inputs attached to that state). For each state a combination function is chosen to combine the causal impacts of other states on this state. It may either use a one common chosen combination functions to aggregate the impacts of multiple states on a given state, or different combination
functions for different states. Though it may be possible to use many combination functions, the logistic sum combination function was identified as a good option for computational cognitive models. Therefore, it was mainly used in all the designed cognitive models though there is a room to further research on new combination functions. Finally, the speed of change of a state is also considered as a parameter. From neurocognitive evidence also it has been found that different speeds occur especially when a state is interacting with external inputs and internal processing. Therefore, in this approach, for each state a speed factor is used to represent how fast a state is changing upon causal impact. Mainly only two speed factors were used, one for states exposed to external environment with low speed values whereas for more internal cognitive states high speed values. This process has been explained in (Treur, 2013, 2016) with more details with possible options for each factor. In summary, the systematic transformation from a conceptual into a numerical representation of the model works as follows (Treur, 2016):

- At each time point \( t \) each state \( Y \) in the model has a real number value in the interval \([0, 1]\), denoted by \( Y(t) \)
- At each time point \( t \) each state \( X \) connected to state \( Y \) has an impact on \( Y \) defined as \( \text{impact}_{X,Y}(t) = \omega_{X,Y} X(t) \) where \( \omega_{X,Y} \) is the weight of the connection from \( X \) to \( Y \)
- The aggregated impact of multiple states \( X_i \) on \( Y \) at \( t \) is determined using a combination function \( c_{Y}(...) \):
  \[
  \text{aggimpact}_{Y}(t) = c_{Y}(\text{impact}_{X_{1},Y}(t), \ldots, \text{impact}_{X_{K},Y}(t)) = c_{Y}(\omega_{X_{1},Y} X_{1}(t), \ldots, \omega_{X_{K},Y} X_{K}(t))
  \]
  where \( X_i \) are the states with connections to state \( Y \)

- The effect of \( \text{aggimpact}_{Y}(t) \) on \( Y \) is exerted over time gradually, depending on speed factor \( \eta_Y \):
  \[
  Y(t + \Delta t) = Y(t) + \eta_Y [\text{aggimpact}_{Y}(t) - Y(t)] \Delta t
  \]
  or
  \[
  \frac{dY(t)}{dt} = \eta_Y [\text{aggimpact}_{Y}(t) - Y(t)]
  \]
- Thus the following difference and differential equation for \( Y \) are obtained:
  Difference equation: \( Y(t + \Delta t) = Y(t) + \eta_Y [c_{Y}(\omega_{X_{1},Y} X_{1}(t), \ldots, \omega_{X_{K},Y} X_{K}(t)) - Y(t)] \Delta t \)
  Differential equation: \( \frac{dY(t)}{dt} = \eta_Y [c_{Y}(\omega_{X_{1},Y} X_{1}(t), \ldots, \omega_{X_{K},Y} X_{K}(t)) - Y(t)] \)
Two particular combination functions are often used: the identity function $\text{id}(\ldots)$, the scaled sum function $\text{ssum}_\lambda(V_1, \ldots, V_k)$, and the advanced logistic sum combination function $\text{alogistic}_{\sigma, \tau}(\ldots)$ (Treur, 2016):

$$
c_i(Y) = \text{id}(V) = V
$$

$$
c_i(V_1, \ldots V_k) = \text{ssum}_\lambda(V_1, \ldots, V_k) = (V_1 + \ldots + V_k) / \lambda
$$

$$
c_i(V_1, \ldots V_k) = \text{alogistic}_{\sigma, \tau}(V_1, \ldots, V_k) = \left( \frac{1}{1 + e^{-\sigma(V_1 + \ldots + V_k - \tau)}} - \frac{1}{1 + e^{\sigma \tau}} \right) (1 + e^{-\sigma \tau})
$$

Here $\sigma$ is a steepness parameter, $\tau$ a threshold parameter, and $\lambda$ is the scaling factor. These three functions are monotonic. The advanced logistic sum combination function has the property that activation levels 0 are mapped to 0 and it keeps values below 1. The first function $\text{id}(\ldots)$ is often used for states $Y$ that have just one impact from another state $X$; then the following difference or differential equation is obtained:

**Difference equation:** $Y(t + \Delta t) = Y(t) + \eta_Y (\omega_{X,Y} X(t) - Y(t)) \Delta t$

**Differential equation:** $\frac{dY(t)}{dt} = \eta_Y (\omega_{X,Y} X(t) - Y(t))$

Having a numerical representation together with dynamic factors, the models were implemented as a computer program (mainly by using the programming language Java). Dynamical models are used to express and understand behaviours of phenomena over time. Nevertheless, having an implemented model it may still not behave as expected. In these processes the states always depends on the effects of related states at the previous time instant. Computational dynamical models are mainly inspired by the causal relations between states or events. The accuracy of a computational dynamical model strongly depends on the values of its parameters (and of course also on the underlying assumptions). Therefore, it is very important to find the most suitable (the best fitted) values for the parameters such that the results of the model relate well to given observations of the real world phenomenon. Parameter estimation is the process of finding the suitable values for the parameters. Parameter estimation methods for computational dynamical systems are different from other well-known methods such as methods for function approximation, linear and non-linear regression, transformations, multivariable differential equations, etc. This is mainly because of not knowing directly a function(s) for behaviour(s) of an output(s). In parameter estimation it is assumed that actual real world data are given for one or more time points for at least one variable (or a state) from the real
world phenomenon in. These real world data are compared with the data from the model for that given particular variable (see Fig. 1).

Often it is not possible to obtain complete data for observed behaviour, and in such a situation at least some (partial) data is needed only for some time points or time intervals. These data can be even in the form of patterns or characteristics. Fig. 1, presents the idea of parameter estimation in a dynamical model. Often the considered phenomena in the real world are quite complex. Therefore the aim is to develop a computational dynamical model that will simulate the behaviour of the actual processes. Once such a model has been designed it is needed to tune the model to the actual characteristics of the humans and the situation in general, which are represented by the parameters of the model. In that way the predicted behaviour of the dynamical model can be made aligning with available observations on behaviours of the actual process for the same input. If there is a mismatch between these two, the parameters of the model have to be adjusted until the difference is acceptable. Mainly two parameter estimation methods are used in this research: one is an analytically driven method and the other one is an improved partial swarm optimisation based method.

Simulations are the last phase, in which one performs experiments with the model to generate traces. Depending on the complexity of the model, the number of its parameters can vary from just a couple to hundreds. These parameters usually represent specific characteristics of the modelled phenomenon. Having identified

![Diagram](image)

**Figure 1:** Parameter estimation for computational dynamical models: In this problem the main process is to compare the observed behaviour from the actual phenomenon with the predicted behaviour from the dynamical model at some time points and if there is a significant difference (the objective should be to make the difference minimum), the model’s behaviour has to be made closer to the observed behaviour by changing the values of the parameters. More specifically, the various parameter estimation methods specify how to quantify the difference and based on that in what way (by an algorithm) the values of the parameters should be adjusted.
the best fitted values for parameters, it is then important to validate the behaviour of
the model mainly through simulations. In general it is possible to find parameters
specifically for each simulation scenario. Nevertheless, this approach is very tedious
and time consuming and more importantly not useful in applications, as it needs to
find a complete set of new values for parameters for each scenario. Therefore, in
this research always more attention was given to find a unique parameter value set
for each model and with a very minimal change to selected parameters that provide
results for different scenarios through simulations. This can be used as a strong
point to justify the correctness of the designed model. By designing a computational
cognitive model based on the latest neuro-cognitive evidence, it is providing valid
simulation results for different scenarios based on one unique parameter value set;
this provides more reliability and confidence on the designed model’s correctness.

1.5. Thesis outline

This thesis is based on collection of conference and journal papers where I
myself have the authorship of all the papers (one paper is not published yet). In all
the papers with more than one author, the different authors have made a comparable
contribution to the paper. Furthermore, there is some overlap between the papers;
nevertheless each paper has its own scope. Following is the list of papers that have
been used in this thesis.

1.5.1. Journal papers

Journal papers included in this thesis:
Thilakarathe, D. J. (2015). Modelling of situation awareness with perception,
attention, and prior and retrospective awareness. Biologically Inspired
Cognitive Architectures, 12, 77–104.
http://doi.org/10.1016/j.bica.2015.04.010

Thilakarathe, D. J., & Treur, J. (2015a). Computational cognitive modelling of
http://doi.org/10.1007/s40708-015-0013-3

http://doi.org/10.1016/j.bica.2015.07.001

1.5.2. Conference papers

Conference papers included in this thesis:
Ham, W. van der, Klein, M. C. A., Tabatabaei, S. A., Thilakarathe, D. J., & Treur,
J. (to appear). Adaptive Methods for a Smart Thermostat to Estimate the
Characteristics of a House Based on Sensor Data with Varying Extent of

http://doi.org/10.1109/WI-IAT.2015.141


http://doi.org/10.1016/j.egypro.2014.06.072

http://doi.org/10.1109/WI-IAT.2014.168

http://doi.org/10.1007/978-3-319-09891-3_42

http://doi.org/10.1007/978-3-319-19066-2_9

1.5.3. Main Parts of this thesis

This thesis consists of five parts, each focusing on different research questions, stated in sub section 1.2 above. Below, each part of the thesis is outlined.

- **Part I: Introduction**
  In the introduction, the main research question within this thesis is introduced: how can dynamic computational cognitive models for human action selection be designed, developed, simulated and applied, and what is the role of awareness and cognitive control in such models? Cognitive, behavioural and affective science related research was explored in the direction of the above theme. This main research question is refined into more atomic research questions, each of which is more manageable and sufficiently specific to explore.

- **Part II: Modelling the Role of Awareness, Emotion, Ownership and Control in Human Action Selection**
  The roles of cognitive processes related to action selection are interesting to study, but they are challenging due to the complexity in human brain. Within
these cognitive processes for action selection awareness, emotion, ownership and control play important roles. This part of the thesis focuses on modelling action preparation and performance by considering its cognitive effects and affects from both a prior and a retrospective perspective relative to the action execution. How action selection and execution contribute to action awareness or vice versa is also considered. Emotions are strongly coupled with most of the other processes related to action selection. The role of emotional awareness and its contribution to and impact on the other cognitive processes is also explored. Furthermore, ownership of action is considered as this is important to differentiation in how far a person attributes an action to him or herself, or to another person. Furthermore, similar to the importance of action selection, cognitive control is also a vital process, especially for behaviour as a healthy socially accepted person. For this purpose the interplay between bottom-up and top-down processes were explored particular to cognitive control in action selection.

- **Part III: Modelling Cognitive Metaphor in Joint Decision Making**
  Making decisions together with others is an essential part of human life, especially in social and professional context. It is interesting to explore what can be learnt from cognitive metaphors for joint decision making in particular. Therefore, the influence of cognitive metaphors on joint decision making has been examined by combining the concept of a joint decision making process and cognitive metaphor together in a computational social agent model. In close relation to this, more in general, understanding new things based on what already has been understood is also essential, which is referred as analogy making. Analogy making is a fundamental process of human cognition. Therefore, this research exploits the Buddhist theory of Five Aggregates to model analogy making. This shows the possibility of adopting different types of literature in cognitive modelling.

- **Part IV: Application of Integrative Computational Models in Complex Real-World Domains: Aviation and Domestic Energy Management**
  As part of the process of designing and developing cognitive models it is important to evaluate and explain how to apply them in real world situations. Two domains were selected for this purpose: aviation and energy management. For the aviation domain situation awareness was simulated through a dynamic cognitive model that was developed. For the energy domain, some research was conducted to evaluate the use of the type of models developed in heat pump related energy management. Therefore, an analytical model for mathematical analysis of smart daily energy management for air to water heat pumps was developed and it was separately validated with some real data too. Having this analytical model (with separate validations)
and a cognitive model for action selection, these two were integrated to provide insights for cognitive driven energy management choices. In addition, in all applications of cognitive models an important challenge is parameter estimation. Therefore, an improved parameter estimation method was developed that is particularly useful for dynamic cognitive models for which not much empirical data can be collected.

- **Part V: Discussion**
  This part summarises the research presented and highlights implications of this work. Furthermore, an outline of some future work is included providing links to new interesting challenges to explore.

### 1.6. Personal contribution to each chapter

In this thesis each chapter represents a published research paper (except Chapter 6), where I myself have the authorship of all the chapters. In our group there has been a longstanding policy that to have an authorship in a research paper each individual should have a comparable contribution (both in doing the research and writing the paper). Therefore, in all the papers with more than one author, each author has made a comparable contribution to the paper and therefore always author names are always in alphabetical order of their surnames. Furthermore, due to this policy I got the opportunity to have single authorship for a few papers although I was supported and guided continuously in each work. The following table summarises my personal contribution to each chapter:

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 2</td>
<td>This chapter is a published journal paper which is an extended version of a conference paper. In the conference paper I was mainly involved in implementing the designed model, simulating four scenarios with estimations for parameter, and writing those parts in the conference paper. This chapter, as being the journal paper, my contribution was extending the literature and incorporating lateral inhibition into the model. Furthermore, I contributed to an extension of the number of simulations from four to eight, together with improving and including new text to the paper with a detailed discussion section.</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>This chapter is also a published journal paper which is an extended version of a conference paper. In the conference paper my main contribution was to explore the literature and to find the cognitive process behind intentional inhibition. Also I was strongly involved in the model design phase introducing effect prediction and impact prediction loops. I also implemented the model and proposed a new analytical approach to parameter estimation. I was involved in simulating two scenarios and writing some parts in the conference paper.</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>This chapter is a published conference paper. My contribution includes exploring cognitive and affective neuroscience literature and to introduce bottom-up and more top-down related processes for emotion generation and incorporating those into the cognitive model. Furthermore, in this work a Java based system was designed and implemented that can be used to automate the implementation process of a cognitive model more easily with minimal time and furthermore, to support manual parameter estimation (especially when many parameters exist) in a more effective manner with XML based inputs. In addition, I was involved in simulating scenarios and writing the parts of the paper.</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>This chapter is a published conference paper for which I am a single author.</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>This work is mainly from my Masters research and prof. Asoka Karunananda guided me for this.</td>
</tr>
<tr>
<td>Chapter 7</td>
<td>This chapter is a published conference paper. My main contribution in this work is to implement the model and to obtain generic values to parameters to simulate many scenarios. Furthermore, I contributed to writing some parts of the paper also.</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>This chapter is a published conference paper where I am a single author.</td>
</tr>
<tr>
<td>Chapter 9</td>
<td>This chapter is a published journal paper which is an extension of two conference papers of which I am a single author.</td>
</tr>
<tr>
<td>Chapter 10</td>
<td>This is a published conference. My contribution for this work is to modelling the indoor temperature when cooling down and temperature maintenance energy usage in an analytical model. Furthermore, I implemented the analytical model and obtained values for temperature maintenance energy usage and temperature increase energy usage. Furthermore, I was involved in writing the paper.</td>
</tr>
<tr>
<td>Chapter 11</td>
<td>This is a published conference paper. My contribution for this work includes exploring literature about heat pump performance and the importance of domestic energy management, parameter estimation for usage of a heat pump, and obtained values for energy usage of the heat pump for the given empirical data set both in over the average outdoor day temperature, and over the days. Furthermore, I was involved in writing the paper.</td>
</tr>
<tr>
<td>Chapter 12</td>
<td>This is a published conference paper. My contribution for this work includes contributing for calculating energy usage based on a degree-days driven approach. I was involved in writing the paper as well.</td>
</tr>
<tr>
<td>Chapter 13</td>
<td>This chapter is a published conference paper. My contribution includes in some parts of combining the cognitive model and the energy model, implementing the designed model and evaluation of the model through simulation experiments. I also was writing some parts of the paper too.</td>
</tr>
</tbody>
</table>
References


[http://doi.org/10.2991/978-94-91216-31-2_3](http://doi.org/10.2991/978-94-91216-31-2_3)


http://doi.org/10.1007/978-3-642-11161-7_20

http://doi.org/10.1523/JNEUROSCI.0924-07.2007

http://doi.org/10.1177/1073858408317417

http://doi.org/10.1016/0165-0173(94)00016-I

http://doi.org/10.1145/375735.375766


http://doi.org/10.4414/swm.2013.13786

http://doi.org/10.1038/nrn2555


http://doi.org/10.1016/j.neuroimage.2011.11.079

http://doi.org/10.1109/TSMC.1982.4308795

http://doi.org/10.3389/fnhum.2012.00179

http://doi.org/10.1007/s10462-009-9094-9


http://doi.org/10.1093/med/9780199837755.001.0001


http://doi.org/10.1007/3-540-49057-4_1


http://doi.org/10.1017/S0048577201393198

http://doi.org/10.1038/nrn2497

http://doi.org/10.1038/nn827


http://doi.org/10.1016/S1364-6613(02)01913-7

http://doi.org/10.1162/jocn.2009.21189


http://dx.doi.org/10.1017/S0140525X07003214


http://doi.org/10.1177/1073858413514136

http://doi.org/10.1109/ICNN.1995.488968

http://doi.org/10.1016/j.cub.2010.03.013


Thera, P. (1972). The Psychological Aspect of Buddhism. BPS.


ResearchGate URL: 


http://doi.org/10.1016/j.actpsy.2012.11.014

http://doi.org/10.1016/j.neuropsychologia.2009.10.026


http://doi.org/10.1016/S1364-6613(97)01070-X

http://doi.org/10.1016/j.neuroimage.2012.06.058
Part II:

Modelling the Role of Awareness, Emotion, Ownership and Control in Human Action Selection

The roles of cognitive processes related to action selection are interesting to study, but they are challenging due to the complexity in human brain. Nevertheless, due to the strong developments in brain research related to experimental equipment and designs, more and more interesting information is discovered. Within these cognitive processes for action selection awareness, emotion, ownership and control play important roles. This part of the thesis focuses on modelling action preparation and performance by considering its cognitive effects and affects from both a prior and a retrospective perspective relative to the action execution. How action selection and execution contribute to action awareness or vice versa is also considered. Emotions are strongly coupled with most of the other processes related to action selection. The role of emotional awareness and its contribution to and impact on the other cognitive processes is also explored. Furthermore, ownership of action is considered as this is important to differentiation in how far a person attributes an action to him or herself, or to another person. Furthermore, similar to the importance of action selection, cognitive control is also a vital process, especially for behaviour as a healthy socially accepted person. For this purpose the interplay between bottom-up and top-down processes were explored particular to cognitive control in action selection.
Chapter 2
Computational Cognitive Modelling of Action Awareness: Prior and Retrospective

Dilhan J. Thilakarathne, Jan Treur
Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands. 
e-mail: d.j.thilakarathne@vu.nl, j.treur@vu.nl

Abstract: This paper presents a computational cognitive model for action awareness focusing on action preparation and performance by considering its cognitive effects and affects from both prior and retrospective form relative to the action execution. How action selection and execution contribute to the awareness or vice versa is a research question and from the findings of brain imaging and recording techniques more information has become available on this. Some evidence leads to a hypothesis that awareness of action selection is not directly causing the action execution (or behaviour) but comes afterward as an effect of unconscious processes of action preparation. In contrast, another hypothesis claims that both predictive and inferential processes related to the action preparation and execution may contribute to the conscious awareness of the action and furthermore, this awareness of an action is a dynamic combination of both prior awareness (through predictive motor control processes) and retrospective awareness (through inferential sense-making processes) relative to the action execution. The proposed model integrates the findings of both conscious and unconscious explanations for both action awareness and ownership and acts as a generic computational cognitive model to explain agent behaviour through the interplay between conscious and unconscious processes. Validation of the proposed model is achieved through simulations on suitable scenarios which are covered with actions that are prepared without being conscious at any point in time, and also with the actions that agent develops prior awareness and/or retrospective awareness. Having selected an interrelated set of scenarios a systematic approach is used to find a suitable but generic parameter value set which is used throughout all the simulations that highlights the strength of the design of this cognitive model.

Keywords: Awareness, Prior, Retrospective, Cognitive Modelling

1 This chapter was published as:
http://doi.org/10.1007/s40708-015-0013-3

Which is an extended work of the following conference paper:
http://doi.org/10.1007/978-3-642-38637-4_7

The names of the authors are ordered alphabetically reflecting the comparable contribution of each author
2.1. Introduction

Humans intuitively feel that their behaviour is an effect of their conscious decisions for certain actions aiming for desired outcomes [1]. However, what exactly is consciousness is a well-known question among many scientists in many disciplines (see [2, 3]); for example, is it just a process in the brain, and if so, how is it composed? With the developments in brain imaging and recording techniques more and more detailed information on various brain processes becomes available, including the conscious awareness of actions. One of the leading hypotheses for action awareness is that humans may prepare for and perform actions without being conscious of these preparation and execution processes. More specifically, the feeling of intention for an action is not causing the behaviour but comes after the action preparation and just before the action execution time [4–8]. It has been found that for certain types of actions the decision to perform it is already made at least hundreds of milliseconds (and even up to 10 seconds) before any awareness state occurs [4, 5, 8]. The brain predicts the outcome of a decision even before the decision reaches awareness and humans’ (illusionary) consciousness seems like an after-effect of a set of unconscious cognitive processes leading to the action [1, 4–8].

The human brain is a complex, intricate, adaptive, dynamic system; it is difficult to unravel it and comprehend its mechanisms (cf. [9–14]). Therefore, given the complexity and contradictions observed in different experiments, alternative hypotheses are also proposed on action awareness (cf. [15–19]). In particular, it is interesting to analyse how an individual is acquiring skills related with a task that he/she is not familiar before (i.e. before something is becoming a habitual task) and how awareness contributes in such a situation [2]. For example, someone who is in the mid age but has no prior experience of cycling, it is an interesting phenomenon to find the experiences gone through in the learning process to develop this skill. Most probably at the beginning of this task it may be really challenging and much attention and awareness are required to perform the task. As a result of that he/she may not be able to change the (visual) focus at all to other environmental cues in order to maintain the balance, keep the ride straight, take turns safely, and pedal up a sloop. Most probably combining what is predicted and what is observed thereby minimizing the error (prediction vs. actual) is the basis for the collected experience in the learning period. Furthermore, how the awareness affects the neural plasticity is an interesting phenomenon too. Although in the learning period a high level of awareness on the action is important, once it becomes a habitual task such awareness may practically become absent: most probably he/she may not pay any attention at all on pedalling, balancing, and keeping straight. Therefore, in daily life experiences predictive and inferential processes for action awareness are important. Nevertheless, when a person learned how to ride a bicycle properly, still he/she may
Moore & Haggard [20] have investigated how predictive or inferential processes of action execution play a role in conscious awareness of an action. They have proposed that awareness of an action is a dynamic combination of both prior awareness (i.e., awareness of the action prediction) and retrospective awareness (i.e., the awareness of the effects of an action), through predictive motor control and inferential sense-making relative to the action execution, respectively (cf. [20, 21]). When a prior awareness state occurs, he/she may become aware of going to perform the action. Having such a prior awareness state still may leave open whether the agent is able to consciously decide to perform or not to perform the action (cf. [1, 4–8]). For example, is still some form of vetoing of the action possible? In principle, the awareness state may play the role of generating a kind of green light for execution of the action. However, equally well the prior awareness state may just play the role of a warning for the agent to be prepared that the action will happen (anyway). As stated in [5, 8] it has been found that for certain types of actions the decision to perform it is already made at least hundreds of milliseconds (and even up to 10 seconds) before any awareness state occurs. These findings may suggest that prior awareness often will have no effect on the decision. But this may strongly depend on the type of action. For example, it will be difficult to believe that the action of buying a car or a house remain unconscious and may not be amendable to vetoing based on awareness states (the Monty Hall Problem is also another good situation for this concern). An awareness state can also develop in retrospect, after the action was performed and this will answer the question: ‘what have I done?’ Such a retrospective awareness state often relates to acknowledging others from and taking responsibility for having performed the action. It may also play an important role in learning as mentioned in the previous example (i.e., by evaluating the obtained effect in a conscious manner leads to improvement of the performance of the action selection).

This paper extends the work published in [22] by refining the neurologically inspired agent model with more realistic simulation results, new scenarios, and a detailed formal specification of the model, together with a more sophisticated parameter estimation methodology. The selected scenarios include a reasonable spectrum of situations in which (a) actions are prepared without being conscious at any point in time, (b) the agent develops prior awareness or retrospective awareness, or both. An example of a schizophrenic patient and an early stage that may lead to a depression situation are also included. As research questions this paper mainly contributes to:
1. How does the internal prediction process shape or contribute to the (prior) awareness of the action?
2. How does the inferential sense making shape or contribute to the (retrospective) awareness of the action execution?
3. How does the awareness contribute to action execution?
4. What is the relation and interplay between conscious and unconscious action formation through action ownership and relevant awareness states?

In addition to awareness states, ownership states for an action are also considered in this paper. They are mainly used as important states in the unconscious action formation process. The specific role of the ownership states (in prior and retrospective form) has been separately discussed in [23]; such a more detailed overview is not included in this paper. The structure of the paper organised with a conceptual basis that includes evidences from cognitive neuroscience which followed with a model description in which a detail explanation of the model together with its mathematical basis and formal specification will be presented. To validate the workings of the proposed model eight scenarios are simulated through a unique parameter value set that estimated using a systematic approach. Finally a discussion will be presented highlighting the usefulness of a model in this nature and future works.

2.2. Action Awareness Viewed Neurologically, Psychologically and Behaviourally

In neurological, psychological, and behavioural literature the notions of awareness and ownership of an action have received much attention. Action ownership is a useful concept which is mainly important to differentiate in how far a person attributes an action to him or herself, or to another person (see [23]). Although in many cases the feeling that you get when you perform something or another person is performing the same action may be similar, it is clearly possible to identify whether the action belongs to you or to someone else. More importantly the information about another person’s behaviour influences your self evaluation and vice versa, which makes humans social beings (cf. [24, 25]). After the discovery of Mirror Neurons such social phenomena including empathy, imitation, and coordination in a social context can be explained more scientifically as a cognitive process [26–28]. Mirror neurons have been mainly identified in two cortical areas: the posterior part of the inferior frontal cortex and the anterior part of the inferior parietal lobule [28]. They have shown strong correlations not only with specific movements, but also with specific goals (or goal directed actions: e.g. reaching for and grasping an object). From the development perspective of human cognition on self and other representations, their interconnection and how those relate to the cognitive processing were highlighted:
‘Over the first several years of life, children acquire knowledge of both objective and subjective aspects of self and others. By 18–24 months of age infants can recognize their own mirror image, a capacity that has been linked to the emergence of self-conscious emotions (e.g. embarrassment [...]). During the preschool years, children simultaneously develop the capacity to represent their own and others’ mental states [...]. This development entails the ability to recognize when self and other perspectives and experiences are shared and thus congruent, and under which circumstances they differ from one another. Interestingly, the development of mental state understanding is functionally related to executive functions [...], suggesting that the prefrontal cortex is implicated in self/other cognitive representations. Indeed, neuroimaging data suggest that theory of mind tasks and executive function tasks share overlapping areas of activation in the medial prefrontal cortex.’ ([24], pp. 527-528)

This separation of self and other is contributing to the ability to recognize ownership of your own action. Research has shown evidence that action prediction (based on sensory information) leads to an action execution with ownership, while when there are problems with action prediction that leads to abnormal states of ownership of that action. For example, when you are tickled by someone, as you have not predicted the action you will experience various sensations due to this sudden action, and the ownership of that action may not be with you (though you do have the body ownership in this situation) (cf. [29]). Chaminade, and co-workers [25] have highlighted this:

‘The neural underpinnings of internal models for motor control have been investigated with human non-invasive neuroimaging techniques (for review see Wolpert and Kawato, 1998). Motor commands that are used by forward models to suppress sensory signals are believed to originate upstream from the primary motor cortex (Voss et al., 2006), though they may also involve premotor areas in the posterior inferior frontal gyrus (Kilner et al., 2007). Actual sensory feedback is used to compute prediction errors for model evaluation and update. When we are tickled by another person (Blakemore et al., 1998) the sensory consequences of its actions are unpredictable, and the lack of predictability leads to a high prediction error associated with increased activity in the secondary somatosensory cortex. This area, located bilaterally in the parietal opercula (Eickhoff et al., 2006), plays a key role in sensorimotor integration (Inoue et al., 2002), and has been involved in the assessment of action ownership (Blakemore and Frith, 2003).’ ([25], pp. 2)

The nature of human actions varies from direct responses on stimuli to actions that takes longer periods to process and react. Here the first types of action are often labelled as automatic or unconscious and the other types as more conscious or intentional [30]. In contrast to action ownership, action awareness is a conscious state. Patients with the Anarchic Hand Syndrome (AHS) (patients with frontal lobe and callosal damage (c.f. [31])) always have some form of ownership and awareness of their action but are not able to control the action [32]. For example, in a cafe just seeing a cup of coffee of an unknown person, for an AHS patient might
be sufficient to reach and grasp it due to the automatic activation of action plans on this habitual task. In normal context this has been trained as a habitual task, but with the (prior) awareness of people who know when he or she should do this. A few more of such AHS examples are grabbing a doorknob or scribbling with a pencil or combing one’s hair [31]. Furthermore, persons suffering from schizophrenia may easily attribute self-generated actions to (real or imaginary) other persons (see [23]). Furthermore, it has been noted that the problem with AHS patients is to control their action while with schizophrenic patient it is a problem with awareness of the action (an AHS patient tries to prevent abnormal behaviour of the alien hand by the good hand after it executed the action) [31]. Through learning with intention and awareness, people have pre-stored actions per stimulus and later without the intention or the awareness the brain will automatically evoke the relevant action which was habitually associated [33]. The frequency and recency of a learned habitual task seems to relate to its probability of getting selected.

The research on action awareness is a challenging task and it is assumed to be that individuals are aware only of the tip of the action iceberg; much further research is necessary to explore and refine the body of knowledge on this (cf. [2, 3, 8, 11, 15, 21]). Nevertheless there are interesting research findings on this. Empirical evidence collected through an experiment setup proposed by Benjamin Libet and his colleagues [8] has challenged the traditional view of human will and has shown that the brain initiates voluntary movements before we are aware of having decided to move. From a cognitive neuroscience perspective human actions are mainly a result of signals getting to motor neurons (motoneuron) in the spinal cord mainly via the primary motor cortex and some of its neighbour areas (e.g. pre motor cortex, supplementary motor cortex (SMA)). Early activation of the primary motor cortex before the agent gets the conscious intention to move (or to act) is called readiness potential and this begins hundreds of milliseconds or even up to 10 seconds before any awareness state occurs [4, 5, 8]. Therefore, it was proposed that conscious will is an illusion and it is too slow to initiate an action, but action formation is due to an unconscious causal chain of processes and just before the action execution we will develop the awareness of the action (not as the cause of the action but as an effect of unconscious processes). John-Dylan Haynes has further improved the Libet experiment setup to advance beyond the shortcomings of the experiment (see [5]) and with his findings again the importance of exploring the tightness of the link between unconscious predictive brain processes and subsequent decisions from a conscious perspective is highlighted:

"An important point that needs to be discussed is to what degree the finding of choice-predictive information supports any causal relationship between brain activity and the conscious will. Such causal links have been demonstrated previously by direct cortical stimulation over parietal and frontal cortex. However, it is unclear if the early
predictive signals are also causally involved in the decision. As for the criterion of temporal precedence, there should be no doubt that our data finally demonstrate that brain activity can predict a decision long before it enters awareness. A different point is the criterion of constant connection. A constant connection would require that the decision could be predicted with 100% accuracy from prior brain activity. Libet’s original experiments were based on averages, so no statistical assessment can be made about the accuracy with which decisions can be predicted. Our prediction of decisions from brain activity is statistically reliable, but far from perfect. The predictive accuracy of around 60% (which is significant, but only 10% above chance) can be improved if the decoding is tailored to each subject. However, even under optimal conditions, this is far from 100% for several reasons. ... ... ... Importantly, a different interpretation could be that the inaccuracy simply reflects the fact that the early neural processes might only be partially predictive of the outcome of the decision. In this view, even full knowledge of the state of activity of populations of neurons in FPC and in the precuneus might not permit the full prediction of a decision. In that case, the signals have the form of a biasing signal that influences the decision to a degree, but additional influences at later time points might still play a role in shaping the decision. The fact that decoding after the decision from motor cortex can be achieved with higher accuracy might point toward the fact that neural signals in BA10 and in PC are not fully predictive in principle. However, the exact topology of clustering of calls with similar tuning preferences in BA10/PC is, to date, unknown, and thus might turn out to be less suitable for fMRI decoding than in motor cortex.’ ([5], pp. 16-17)

With the concerns highlighted in the above quote (for more criticisms on this hypothesis see [34]), though the awareness state emerges just before the action execution it is not yet clear whether there is not at all an impact on action execution from this subjective awareness. One of the issues that have turned out to play an important role both in the execution decisions for an action, and in its attribution, is the prediction of the (expected) effects of the action, based on internal simulation starting from the preparation of the action [35, 36]. If these predicted effects are satisfactory, this may entail a ‘go’ decision for the execution of the action, thus exerting control over action execution. In contrast, less satisfactory predicted effects may lead to a ‘no go’ decision (cf. [37–39]). Predicted action effects also play an important role in attribution of the action to an agent after it has been performed. In neurological research it has been found that poor predictive capabilities are a basis for false attributions of actions, for example, for patients suffering from schizophrenia [31, 40, 41]. In addition to the predictive effects, the sensation of the actual effect (after executing the action) also has been noted as important in action formation research [15, 17, 21, 30, 36]. In literature it has been reported that the predicted sensory effect and the sensed actual effect are integrated with each other as a basis for proper attribution of the action [20, 40, 41]. Another element, put forward in [20], is the distinction between action awareness based on prediction
(prior to execution), and action awareness based on inference after execution of the action (in retrospect):

‘Our results suggest that both predictive and inferential processes contribute to the conscious awareness of operant action. The relative contribution of each of these processes seems to be context dependent. When we can predict the consequences of our actions, as in a high action-effect contingency block, the awareness of action reflects these predictions. This would provide us with a predictive sense of our own agency. In addition, our results show clear evidence that inferential processes also influence the conscious awareness of operant action. ... ... ... The interaction between predictive and inferential processes is of particular interest. ... ... ... The time course over which information about action is built up may be an important clue to this interaction. ... ... ... Sensory feedback provides more precise evidence about actions and their effects. This evidence becomes available only after a short sensory delay, but can then be transferred to memory. Thus, reliable and enduring sensory evidence replaces short-lived predictive estimates. We suggest that awareness of action therefore switches from a predictive to an inferential source as the action itself occurs, and as sensory information becomes available.’ ([20], pp. 142-143)

With these evidences they have suggested that awareness of an action is a dynamic combination of both prior awareness (i.e., awareness of the action effect prediction) and retrospective awareness (i.e., the awareness of the effects of an action) through predictive motor control and inferential sense-making relative to the action execution, respectively [15, 21, 30, 36] (c.f. [19]). Furthermore, Haggard and co-workers presented a new phenomenon called intentional binding: when a voluntary action produces (with the awareness and intention) the temporal (subjective) gap between the action and its perceived sensory outcome is less when the awareness is pre-existing but it is high when the awareness does not involve this [15]. This phenomenon has been argued to be an effect of either prior awareness or retrospective awareness with different experiment setups. To investigate the relation with prior awareness, transcranial magnetic stimulation (TMS) was randomly applied over the motor cortex and the entailed disruption of awareness observed (here through the intention in this setup). A significantly weakened intentional binding has been observed; therefore, it may be useful to highlight the necessity of prior awareness (see [19]). Similarly, there are experiments to analyses the influence from retrospective awareness selecting some tasks where prediction of the action outcome is difficult (or unpredictable) but there is a ‘tone’ after the action execution [20] and it has been observed that retrospective processes play a role when prior predictive processes are absent (or when prediction was minimal). In addition to the mentioned roles in prior and retrospective effects of intentional binding there are some evidences for its neural basis also:

‘Moore and colleagues investigated the contribution of two specific target sites: the pre-supplementary motor area (pre-SMA) and primary motor cortex (M1). The pre-SMA is
involves in higher-order cognitive aspects of self-generated action (Picard & Strick, 2001) and with the conscious experience of intending to act (Fried et al., 1991). In this sense it is likely to support predictive contributions to intentional binding. On the other hand, M1 processes signals that are involved in actual motor execution, signals that the authors suggest are required to support inferences of agency. It was found that only stimulation of pre-SMA led to a significant reduction in intentional binding. Stimulation of M1 marginally reduced intentional binding, but this effect was not significant. The authors therefore concluded that pre-SMA is likely to play a key role in intentional binding.' ([19], pp. 5)

Inhibition and suppressive mechanisms may also be as important as the excitation mechanisms in cognitive control (though some different viewpoints are also put forward (see [42])). By using Gamma-aminobutyric acid (GABA), neurons are performing inhibition at synaptic, circuit, and systems levels (cf. [42]). Furthermore, various inhibition types in neuroscience and psychology have been discussed in [42]. Inhibition activates in automatic (e.g., lateral inhibition: if a particular representation accumulates more evidences, that will suppress its fellow representations) and voluntary (e.g., suppression of an irrelevant response, stimulus, or memory; in intentionally) manners. Another peculiar aspect that has been observed is that within the process of co-occurrence of predicted effects and sensed actual effects the predicted effect suppresses the sensed actual effect [29, 43, 44]. Moreover, it has been put forward that the predicted effect and the sensed actual effect are not simply compared or matched, as claimed in the so-called ‘comparator model’ in earlier literature such as [31, 35, 45], but in fact are added to each other in some integration process [20, 40, 41].

Though these evidences are facilitating an adequate level of information to comprehend a theoretical cognitive system in the form of a model, it is further required to confirm these findings; this may have different variants to be explored in future research.

2.3. Description of the Cognitive Computational Model

Having discussed the evidence on awareness (a person’s subjective experience) and ownership (in how far does a person attribute an action to him or herself or to another person) in Section 2, this section presents a computational agent model. This model will be used in agent driven applications where the awareness is paramount (or necessary) for decision making and justifications of actions through communication. More specifically, this model can provide interesting and important input for problem domains concerning performing or learning specific healthy behaviours or lifestyles. In such domains having an idea about the extent of the awareness of decisions concerning health or lifestyle is important and the model may provide the fundamentals for applications in these domains (with further
refinements and customization where needed). Furthermore, this model may be useful in medical domains where people can analyse and compare the phenomenal effects of certain scenarios through various simulations and further to compare and contrast different hypothesis in theoretical manner to conduct more complex experiments. Also computational simulations have become a promising approach to analyse the emergence in complex systems (e.g., in the aviation domain, or in social science) and a model like this may provide more realistic results especially when human cognitive aspects are necessary in such simulations (e.g., simulating situation awareness in air traffic control, including the human cognitive aspects).

An overview of the postulated cognitive agent model is presented in Figure 1 and its abbreviation details can be found in Table 1. The model is a refined version of a previous model presented in [22] but improving the action preparation process and simulation results. In this model awareness states are taken specific for a given action $a$, effect $b$, context $c$, and stimulus $s$. When the context $c$ is self, an awareness state for $a$, $b$, $c$, and $s$ indicates self-attribution awareness, whereas for context $c$ an observed agent B, it indicates awareness of attribution of the action to B. Specific attention is given to ‘self’ than ‘other’ in simulations of this paper (for ‘other’ see

---

**Fig. 1:** Overview of the computational cognitive agent model. Here an arrow $\rightarrow$ represents a direct activation to state B from state A, an arrow $\rightarrow$ represents a direct suppression to state B from state A, an arrow $\cdots\rightarrow$ represents a suppression to all the complements of ‘$i$th’ state on B, from state $A_i$ (where ‘$i$’ presents an instance of a particular state), and $\star\rightarrow$ represents a direct suppression to all parallel forms of that state.
Table 1: Nomenclature for Fig. 1

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WS(W)</td>
<td>world state $W$ ($W$ can be either: context $c_s$, stimulus $s_s$, or effect $b_i$)</td>
</tr>
<tr>
<td>SS(W)</td>
<td>sensor state for $W$ ($W$ can be either: context $c_s$, stimulus $s_s$, or effect $b_i$)</td>
</tr>
<tr>
<td>SR(W)</td>
<td>sensory representation of $W$ ($W$ can be either: context $c_s$, stimulus $s_s$, or effect $b_i$)</td>
</tr>
<tr>
<td>PA(a)</td>
<td>preparation for action $a_i$</td>
</tr>
<tr>
<td>F($b_i^+$)</td>
<td>feeling for action $a_i$ after as-if loop or action execution</td>
</tr>
<tr>
<td>EA(a)</td>
<td>execution of action $a_i$</td>
</tr>
<tr>
<td>PO($a_i, b_i, c_i, s_i$)</td>
<td>prior ownership state for action $a_i$ with $b_i$, $c_i$, and $s_i$</td>
</tr>
<tr>
<td>RO($a_i, b_i, c_i, s_i$)</td>
<td>retrospective ownership state for $a_i$ with $b_i$, $c_i$, and $s_i$</td>
</tr>
<tr>
<td>PAwr($a_i, b_i, c_i, s_i$)</td>
<td>prior-awareness state for action $a_i$ with $b_i$, $c_i$, and $s_i$</td>
</tr>
<tr>
<td>RAwr($a_i, b_i, c_i, s_i$)</td>
<td>retrospective-awareness state for action $a_i$ with $b_i$, $c_i$, and $s_i$</td>
</tr>
<tr>
<td>EO($a_i, b_i, c_i, s_i$)</td>
<td>communication of ownership and awareness of $a_i$ with $b_i$, $c_i$, and $s_i$</td>
</tr>
</tbody>
</table>

Furthermore, causal relationships in the model are based on the neurological literature presented in the Section 2; they do not take specific neurons into consideration but use more abstracted cognitive or mental states for the design of the model (through an interlevel relation between the neurological level and the cognitive/affective mental modelling level). The model uses three world states (WS) as inputs for:

- stimulus $s$: WS($s$)
- context $c$: WS($c$)
- effect $b$: WS($b$)

The stimulus $s$ represents any internal (bodily, e.g. self-generated facial expression) or external change that may lead to an action execution. Context $c$ represents additional information perceived to improve the process of action selection. The context $c$ can be differentiated as ‘self’ and ‘other’ (self-other distinction). Effects of mirroring can be modelled when $c$ is ‘other’ (see [22, 23]). The effect $b$ represents the effects of the execution of an action $a$.

The input world states WS($s$), WS($c$), and WS($b$) lead to sensor states SS($s$), SS($c$), and SS($b$), and subsequently to sensory representation states SR($s$), SR($c$), and SR($b$), respectively. The unconscious causality of action formation has been modelled as explained in [46]: by combining Damasio’s as-if body loop (see [37–39]) and James’s body loop (see [47]) hypotheses. The body loop has been mapped in this model by the following causal relationships:
Damasio extended the body loop concept and argued that the cognitive process of action selection is due to an effect of an internal simulation process prior to the execution of an action [37–39]. The brain will evaluate the effect of each relevant action option (i.e. \(PA(a_i)\)) by comparing the feelings associated to each individual valuated effects (without actually executing them through the body loop). The simulated option that has the strongest valuated feeling performs as a GO signal through the body loop and else are NO-GO options. The as-if body loop consists of:

\[
\text{sensory representation} \rightarrow \text{preparation for bodily changes} \rightarrow \text{felt emotion}
\]

In this model this is represented by the as-if body loop as follows:

\[
PA(a) \rightarrow SR(b) \rightarrow F(b)
\]

The as-if body loop and the body loop demonstrate the working of predicted sensory effects and sensed actual effects, respectively, as highlighted in the Section 2. These processes are mainly considered to be unconscious processes involving multiple options for responses evaluated in parallel, to determine an adequate response associated to a stimulus (cf. [33]). Through this parallel internal action simulation mechanism the agent will not select a random option but the one which has the strongest valuated feeling. Therefore, depending on the weight values attached to each option at that particular moment, the model will show different behaviours in simulations. This purely unconscious mechanism may be interrupted by the effects of awareness (which will be explained later) to select something different by adding some bias to the mentioned process. Being a cyclic process, the effects of an injected bias may have the ability to compete with other options to finally provide a GO signal.

In Fig. 1, state labels are attached with subscript letters \(k\) and \(i\), which indicate, for example, the \(k^{th}\) instance for a stimulus \(s\) (e.g., \(WS(s_k)\)) for a given \(s_k\) stimulus and the \(i^{th}\) option for an action \(a\) (e.g., \(PA(a_i)\)). Therefore, through this model it is possible to have multiple action options either through a single stimulus or from multiple stimuli, depending on the specific model instance.

Each \(PA(a_i)\) state is affected by its associated feeling through the as-if body loop. Moreover, each \(PA(a_i)\) state suppresses its complementary options \(PA(a_j)\) for \(j \neq i\) (as shown in dotted looped red arrow in Fig. 1) proportional to the accumulated strength of that option. This behaviour is in line with the explanation for the lateral inhibition in [42] and will contribute to further strengthen the action selection process. Therefore, naturally the strongest internally satisfied option (which is exceeding a threshold value) will become selected as a result of the unconscious action selection process as explained earlier. The feeling state in this model can be either a positive feeling \(F(b_i^+))\) or a negative feeling \(F(b_i^-))\). A given stimulus \(s_k\) may
trigger multiple preparation options in parallel and those might have different
associated feelings, also in parallel (note that when a feeling state’s activation level
is ‘0’ it is assumed to be a case of no feeling).

Prior ownership states have been integrated with the above mentioned processes
and mainly they get affected from sensory representation states SR(s_k), action
preparation state PA(a_i), and feeling states F(b_i). Also the PO(a_i, b_i, c_b, s_k) states
affect prior awareness states PAwr(a_i, b_i, c_b, s_k), retrospective ownership
states RO(a_i, b_i, c_b, s_k), action execution states EA(a_i), and sensory representation
states SR(b_i) of effects b_i. Having a direct link between SR(s_k) to PO(a_i, b_i, c_b, s_k)
facilitates the embedding of salient features of the input to the ownership and
therefore the agent will be able to relate the input and output (together with RO(a_i,
b_i, c_b, s_k)). Furthermore, the link from SR(c_k) to PO(a_i, b_i, c_b, s_k) facilitates the
necessary behaviour of mirror neurons when c is other (for more details see [23]).
The state PO(a_i, b_i, c_b, s_k) has a suppressive effect on SR(b_i); this provides the
mechanism by which the predicted effect suppresses the sensed actual effect (see
[29, 43, 44]). Similar to prior ownership, once an action is executed retrospective
ownership will develop. The retrospective ownership state RO(a_i, b_i, c_b, s_k) is
affected by the prior ownership state PO(a_i, b_i, c_b, s_k), SR(c_k), F(b_i), and EA(a_i).
Furthermore, RO(a_i, b_i, c_b, s_k) activation has effects on the states RAwr(a_i, b_i, c_b, s_k),
PO(a_i, b_i, c_b, s_k), and EO(a_i, b_i, c_b, s_k). As RO is affected by EA(a_i) and F(b_i) this
provides the cognitive behaviour of retrospective effects as differentiated from prior
behaviour. For more details on ownership states of this model, see [23]. Once
RO(a_i, b_i, c_b, s_k) developed it has a suppressive effect on PO(a_i, b_i, c_b, s_k) and
through this also it is contributing to the cognitive shift from predictive to
inferential.

For each ownership state an associated awareness state may (or may not)
emerge. Awareness states play a higher order cognitive role. The direct links from
ownership and feeling to awareness state realise bottom-up activation. Conversely,
the effects of awareness states on other states realise top-down activation, which is
considered to be a conscious or intended process. Therefore in the presented model
PAwr(a_i, b_i, c_b, s_k) is only affected by PO(a_i, b_i, c_b, s_k) and F(b_i). This is useful to
model the idea of Benjamin Libet and others: brains initiate voluntary movements
before we are aware of having decided to move (also see the simulations in Section
4). Moreover PAwr(a_i, b_i, c_b, s_k) affects PA(a_i) and EA(a_i) and this is reflects the
idea of Haggard and co-workers: there may be an impact from this subjective
awareness state on action execution. By this PAwr(a_i, b_i, c_b, s_k) to PA(a_i) link the
agent can inject some bias to the current unconscious process through awareness.
This may strengthen a weaker action option and improve the predictive feeling of
that option (which may lead to getting it executed). Furthermore, in this model
PAwr(a_i, b_i, c_b, s_k) can also directly strengthen the action execution state. Both the
prior ownership states and the prior awareness states are associated to the predictive aspects of the system. In contrast, retrospective awareness states are associated to the inferential aspects, as highlighted in Section 2: awareness of an action is a dynamic combination of both prior awareness and retrospective awareness through predictive motor control and inferential sense-making relative to the action execution. Once the RAwr(\(a_i, b_i, c_k, s_k\)) state is activated it has a suppressive effect on PAwr(\(a_i, b_i, c_k, s_k\)) and due to this PAwr(\(a_i, b_i, c_k, s_k\)) will weaken and RAwr(\(a_i, b_i, c_k, s_k\)) will be dominant after the action execution. Finally, acknowledging of ownership and awareness of an action is modelled by the connection from the RO(\(a_i, b_i, c_k, s_k\)) and RAwr(\(a_i, b_i, c_k, s_k\)) to the EO(\(a_i, b_i, c_k, s_k\)). Once the state EO(\(a_i, b_i, c_k, s_k\)) has become activated, it has a suppressive effect on SR(\(b_i\)) so that this will allow to stop the inferential sense-making process.

In addition to the above mentioned connections a few more suppressive connections are available, which are shown in orange arrows in Figure 1. These connections are mainly for purposes of having an appropriate scenario. Once a stimulus \(s\) and context \(c\) are activated the agent starts to activate the internal processes as mentioned above and once the agent performed the action, a mechanism is assumed that stops the stimuli as an action effect: the agent has performed the task and due to that environment has changed. Therefore these orange connections: EA(\(a_i\)) to WS(\(s_k\)), EO(\(a_i, b_i, c_k, s_i\)) to WS(\(c_k\)), and EO(\(a_i, b_i, c_k, s_k\)) to WS(\(b_i\)) have been included to stop the input stimulus \(s_k\), input context \(c_k\), and the effect \(b_i\) of action \(a_i\). Furthermore, having two inputs (i.e., \(s_1, s_2, c_1,\) and \(c_2\)) if only one action is executed (let’s say \(i=1\): EA(\(a_{i1}\)) and EO(\(a_{i1}, b_{i1}, c_{i1}, s_{i1}\))) then it will be assumed that the executed action will suppress all the inputs, for the purpose of an appropriate scenario.

The following is a brief summary of the agent’s internal causality when given stimulus \(s_k\) and context \(c_k\) as inputs:

1. **action effect prediction**: sensory representation of effect \(b_i\) is affected by preparation of an action \(a_i\)
2. **preparation for action** \(a_i\) is affected by sensory representation of \(s_k\), prior-awareness, feeling of effect prediction of action \(a_i\), and the complements of current preparation for action \(a_i\)
3. a **prior ownership state** is triggered based on preparation for action \(a_i\), predicted effects \(b_i\) of \(a_i\), stimulus \(s_k\), retrospective ownership and context \(c_k\)
4. a **prior awareness state** is activated based on feeling of the predicted effect, prior ownership and retrospective awareness
5. **execution of action** \(a_i\) is affected by prior-awareness, prior ownership and preparation for action \(a_i\)
6. a prior ownership state and prior awareness state exert control over the execution of a prepared action (go/no-go decision, vetoing)

7. suppression of the sensory representation of effect $b_i$ by both prior-ownership and communication of ownership and awareness

8. suppression of the prior ownership state when the retrospective ownership state is developed

9. suppression of the prior awareness state when the retrospective awareness state is developed

10. a retrospective ownership state is activated based on co-occurrence of predicted action effects and action effects sensed afterwards

11. a retrospective awareness state is activated based on action effects sensed by execution of action $a_i$, retrospective ownership, and prior-awareness

12. a retrospective ownership state and retrospective awareness are internal states that also can lead to acknowledging authorship of the action (individually), for example, in a social context

13. execution of an action $a_i$ affects the stimulus $s_k$ in the world

14. communication of ownership and awareness affects both context $c_k$ in the world and effect $b_i$ of action $a_i$ in the world

### 2.3.1. Dynamics of the model

Connections between state properties (the arrows in Figure 1) have weights $\omega_k$, as indicated in Table 2. In this table a weight $\omega_k$ has a value between -1 and +1 and may depend on the specific context $c_k$, stimulus $s_k$, action $a_i$ and/or effect $b_i$ involved. By varying these connection strengths, different possibilities for the repertoire offered by the model can be realised and can be aligned with the considered scenario and behaviour. Usually weights are assumed to be nonnegative, except for the inhibiting or suppressive connections. The behaviour of the model (through simulations) depends on the values of each of these weights (together with other parameters). Determining proper values for these parameters is a non trivial task.

In this table the column LP refers to the (temporally) Local Properties: LP1 to LP17 (c.f. [48]) and that specifies the update dynamics of the activation value of the ‘to state’ based on the activation levels of the ‘from states’. For the dynamics of each local property a LEADSTO formalisation is used, which has been shown to be an appropriate approach to model dynamic behaviours of computational cognitive models [48]. LEADSTO is a hybrid modelling language in which a dynamic property or temporal causal relation $a \rightarrow b$ denotes that when a state property $a$ (or conjunction thereof) occurs, then after a certain time delay, state property $b$ will occur (see [48] for the relevance and benefits of LEADSTO in dynamic models). LEADSTO can be compared to Linear Temporal Logic, but differs in the sense that
Table 2: Overview of the connections and their weights. Here the red colour $\omega_k$ indicates negative weights

<table>
<thead>
<tr>
<th>from states</th>
<th>to state</th>
<th>weights</th>
<th>LP</th>
</tr>
</thead>
<tbody>
<tr>
<td>EA($a_i$)</td>
<td>WS($s_1$)</td>
<td>$\omega_1$</td>
<td>LP1</td>
</tr>
<tr>
<td>EO($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>WS($c_2$)</td>
<td>$\omega_2$</td>
<td>LP2</td>
</tr>
<tr>
<td>EA($a_i$), EO($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>WS($b$)</td>
<td>$\omega_3$, $\omega_4$</td>
<td>LP3</td>
</tr>
<tr>
<td>WS($s_2$)</td>
<td>SS($s_1$)</td>
<td>$\omega_5$</td>
<td>LP4</td>
</tr>
<tr>
<td>WS($c_2$)</td>
<td>SS($c_2$)</td>
<td>$\omega_6$</td>
<td>LP5</td>
</tr>
<tr>
<td>WS($b$)</td>
<td>SS($b$)</td>
<td>$\omega_7$</td>
<td>LP6</td>
</tr>
<tr>
<td>SS($s_2$)</td>
<td>SR($s_1$)</td>
<td>$\omega_8$</td>
<td>LP7</td>
</tr>
<tr>
<td>SS($c_2$)</td>
<td>SR($c_1$)</td>
<td>$\omega_9$</td>
<td>LP8</td>
</tr>
<tr>
<td>PA($a_i$), SS($b$), PO($a_i$, $b_i$, $c_i$, $s_k$), EO($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>SR($b$)</td>
<td>$\omega_{10}$, $\omega_{11}$, $\omega_{12}$, $\omega_{13}$</td>
<td>LP9</td>
</tr>
<tr>
<td>F($b$), SR($s$), PAwr($a_i$, $b_i$, $c_i$, $s_k$), PA($a_i$) ($j \neq i$)</td>
<td>PA($a_i$)</td>
<td>$\omega_{14}$, $\omega_{15}$, $\omega_{16}$, $\omega_{17}$</td>
<td>LP10</td>
</tr>
<tr>
<td>SR($b$)</td>
<td>F($b$)</td>
<td>$\omega_{18}$</td>
<td>LP11</td>
</tr>
<tr>
<td>SR($s$), SR($c$), PA($a_i$), F($b$), RO($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>PO($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>$\omega_{19}$, $\omega_{20}$, $\omega_{21}$, $\omega_{22}$, $\omega_{23}$</td>
<td>LP12</td>
</tr>
<tr>
<td>PO($a_i$, $b_i$, $c_i$, $s_k$), F($b$), RAwr($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>PAwr($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>$\omega_{24}$, $\omega_{25}$, $\omega_{26}$</td>
<td>LP13</td>
</tr>
<tr>
<td>PA($a_i$), PO($a_i$, $b_i$, $c_i$, $s_k$), PAwr($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>EA($a_i$)</td>
<td>$\omega_{27}$, $\omega_{28}$, $\omega_{29}$</td>
<td>LP14</td>
</tr>
<tr>
<td>EA($a_i$), F($b$), PO($a_i$, $b_i$, $c_i$, $s_k$), SR($c$)</td>
<td>RO($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>$\omega_{30}$, $\omega_{31}$, $\omega_{32}$, $\omega_{33}$</td>
<td>LP15</td>
</tr>
<tr>
<td>EA($a_i$), RO($a_i$, $b_i$, $c_i$, $s_k$), F($b$), PAwr($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>RAwr($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>$\omega_{34}$, $\omega_{35}$, $\omega_{36}$, $\omega_{37}$</td>
<td>LP16</td>
</tr>
<tr>
<td>RO($a_i$, $b_i$, $c_i$, $s_k$), RAwr($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>EO($a_i$, $b_i$, $c_i$, $s_k$)</td>
<td>$\omega_{38}$, $\omega_{39}$</td>
<td>LP17</td>
</tr>
</tbody>
</table>

Predicate logical state expressions can be used and also real numbers in them. The traces generated for LEADSTO can be seen as continuous time models satisfying the finite variability property: between any two time points there are only a finite number of state changes.

The time delay defined in LEADSTO is taken as a uniform time step $\Delta t$ here. Table 3 below summarises the formalisation of local properties both in LEADSTO format and in differential equation format. This is used as the formalization of the computational form of the cognitive model described. During the processing, each state property has a strength represented by a real number between 0 and 1 through variables $V$ (with subscripts) that run over these values. In dynamic property specifications, this is added as a last argument in the state property expressions. This representation is considered only for the LEADSTO based formalisation. Therefore, the unary predicate representation of each state in the Table 2 was extended to a binary predicate representation by including the state strength $V_n$ (e.g., $EA(a_i)$ to $EA(a_i, V_n)$). Furthermore, the temporal step in the original LEADSTO formalisation (see [48]) is used as a uniform time step $\Delta t$ in this paper.
Table 3: Specification of Local Properties in the hybrid language LEADSTO and in differential equation format. Here \( Y_{i,k} \) means \((a_i, b_i, c_i, s_i)\) and for each LP the first representation is in the LEADSTO and that will be followed by the differential equation format.

<table>
<thead>
<tr>
<th>LP</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP1</td>
<td>( \text{WS}(s_k, V_2) \Rightarrow \text{WS}(s_k, V_2 + \gamma [f(Y_{i,k}) - V_2] \Delta t) )</td>
</tr>
<tr>
<td>LP2</td>
<td>( \frac{d\text{WS}(s_k)}{dt} = \gamma f\left(\omega_1, \text{WS}(s_k) - \text{WS}(s_k)\right) )</td>
</tr>
<tr>
<td>LP3</td>
<td>( \text{EO}(Y_{i,k}, V_1) \Rightarrow \text{WS}(c_k, V_2 + \gamma [f(Y_{i,k}) - V_2] \Delta t) )</td>
</tr>
<tr>
<td>LP4</td>
<td>( \frac{d\text{WS}(s_k)}{dt} = \gamma f\left(\omega_2, \text{EO}(Y_{i,k}) - \text{WS}(c_k)\right) )</td>
</tr>
<tr>
<td>LP5</td>
<td>( \text{EO}(Y_{i,k}, V_1) \Rightarrow \text{WS}(c_k, V_2 + \gamma [f(Y_{i,k}) - V_2] \Delta t) )</td>
</tr>
<tr>
<td>LP6</td>
<td>( \frac{d\text{WS}(s_k)}{dt} = \gamma f\left(\omega_3, \text{EO}(Y_{i,k}) - \text{WS}(c_k)\right) )</td>
</tr>
<tr>
<td>LP7</td>
<td>( \frac{d\text{WS}(s_k)}{dt} = \gamma f\left(\omega_4, \text{EO}(Y_{i,k}) - \text{WS}(c_k)\right) )</td>
</tr>
<tr>
<td>LP8</td>
<td>( \frac{d\text{WS}(s_k)}{dt} = \gamma f\left(\omega_5, \text{EO}(Y_{i,k}) - \text{WS}(c_k)\right) )</td>
</tr>
<tr>
<td>LP9</td>
<td>( \frac{d\text{WS}(s_k)}{dt} = \gamma f\left(\omega_6, \text{EO}(Y_{i,k}) - \text{WS}(c_k)\right) )</td>
</tr>
<tr>
<td>LP10</td>
<td>( \frac{d\text{WS}(s_k)}{dt} = \gamma f\left(\omega_7, \text{EO}(Y_{i,k}) - \text{WS}(c_k)\right) )</td>
</tr>
<tr>
<td>LP11</td>
<td>( \frac{d\text{WS}(s_k)}{dt} = \gamma f\left(\omega_8, \text{EO}(Y_{i,k}) - \text{WS}(c_k)\right) )</td>
</tr>
<tr>
<td>LP12</td>
<td>( \frac{d\text{WS}(s_k)}{dt} = \gamma f\left(\omega_9, \text{EO}(Y_{i,k}) - \text{WS}(c_k)\right) )</td>
</tr>
</tbody>
</table>
As an example let’s consider LP9 according to the LEADSTO formalisation:

LP9 Sensory representation for an effect bi state

If the preparation state for action ai has level Vi
and the sensor state for effect bi has level V2
and the prior self-ownership of action ai for bi, ck, and sk has level V3
and the communication of ownership and awareness of ai for bi, ck, and sk has level V4
then after time duration Δt the state sensory representation for an effect bi will have the level:

\[ V5 + \gamma \left[ f(\omega_{10}V1, \omega_{11}V2, \omega_{12}V3, \omega_{13}V4) - V5 \right] \Delta t \]
This expresses that after time duration $\Delta t$ the value for the sensory representation $SR(b_i)$ of effect $b_i$ has changed from $V_5$ into

$$V_5 + \gamma \left[ f(\omega_{10}V_1, \omega_{11}V_2, \omega_{12}V_3, \omega_{13}V_4) - V_5 \right] \Delta t$$

This means that

$$\Delta SR(b_i) = \gamma \left[ f(\omega_{10}V_1, \omega_{11}V_2, \omega_{12}V_3, \omega_{13}V_4) - V_5 \right] \Delta t$$

or

$$\Delta SR(b_i) = \gamma \left[ f(\omega_{10}PA(a_i), \omega_{11}SS(b_i), \omega_{12}PO(Y_{i,k}), \omega_{13}EO(Y_{i,k})) - SR(b_i) \right] \Delta t$$

This expression in difference equation format can be rewritten into differential equation format:

$$\frac{dSR(b_i)(t)}{dt} = \gamma \left[ f \left( \omega_{10}PA(a_i)(t), \omega_{11}SS(b_i)(t), \omega_{12}PO(Y_{i,k})(t), \omega_{13}EO(Y_{i,k})(t) \right) - SR(b_i)(t) \right]$$

The same formalisation but specifically for the rate of activation change can be presented in its differential form as found in Table 3 under LP 9 second row. Here, $f$ is a function for which different choices can be made. The function $f$ should be a combination function (when a given state has only a single input the identity function $f(W) = W$ is also usable though it is less configurable). For the simulations the combination function $f$ is based on a continuous logistic threshold function $g(\sigma, \tau, X)$ is used as in the equations (1) and (2).

$$f(\omega_{10}y_1, \omega_{20}y_2, \ldots) = g(\sigma, \tau, \sum_i \omega_{ji}y_j)$$

with

$$g(\sigma, \tau, X) = \left( \frac{1}{1 + e^{-\sigma(X-\tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{\sigma\tau})$$

$$\text{when } x > 0 \quad \text{(1)}$$

$$g(\sigma, \tau, X) = 0; \quad \text{when } x \leq 0 \quad \text{(2)}$$

In the above equations, $\sigma$ is the steepness and $\tau$ the threshold; these are configuration parameters that change the shape of the curve and its midpoint on the X-axis. Activation of a state depends on multiple other states that are directly attached to it; therefore incoming activation levels from other states are combined to some aggregated input and perform the activation according to a specification as in LP9 above, or in an alternative differential equation format, as in equation (3) (where $g$ is the logistic function specified in the equations (1) and (2), and $y_i$ is the activation level of state $i$).
Parameter $\gamma$ is an update speed factor, indicating the speed by which an activation level is updated upon received input from other states. In this model two speed factor values are used: one for the internal states (states which are inside the dotted box in Figure 1), and the other for the external states: $WS(s_k)$, $WS(ck)$, $WS(b_i)$, $SS(sk)$, $SS(ck)$, $SS(b_i)$, $EA(ai)$, and $EO(ai, bi, ck, sk)$. The internal states’ speed factor is higher than the external states (adhering to the phenomenon that brain neurons are activating much faster than sensor and effector organs).

To obtain a computational specification for temporal simulation of each state, a difference equation is used in the form of equation (4).

$$\frac{dy_i}{dt} = \gamma_i \left[ g\left(\sigma_i, \tau_i, \sum_j \omega_{ji} y_j\right) - y_i \right]$$ (3)

By having different values for each parameter (i.e., for weight values $\omega_i$, time step size $\Delta t$, slow and fast speed factors $\gamma$, steepness $\sigma_i$, threshold $\tau_i$) the agent can facilitate a wide variety of behaviours. Each LP in Table 3 is represented in a computational form (by using the JAVA language) and the dynamics of the system is achieved through evaluating the causality effects through a set of difference equations as in equation (4). For each discrete time step $\Delta t$ the behaviour of each state is calculated the emergence of the behaviour is traced with a identified parameter value set. From a mathematical point of view, the dynamics of the model is (numerically) solving the differential equations of LP1 to LP17 by assuming that at time $t=0$, $WS(s)$ and $WS(c)$ holds value 1 as state level. Some of these behaviours are presented as simulations in Section 4.

**2.4. Simulation Results**

In this section simulation experiments for a number of example scenarios are discussed, which all involve the occurrence of a preparation state for an action $ai$, triggered by some stimulus $sk$ and context $ck$. These scenarios relate to phenomena in the literature, as discussed in Section 1 and 2. They have been generated based on the specification presented in Section 3. Eight scenarios have been simulated to highlight the different possible behaviours of the model and among those 3 scenarios are new, whereas the other 5 were selected from the previous work in [22] but those behaviours have been significantly improved in the newer versions presented here. Furthermore; for the scope of this paper only $c$ is ‘self’ situations are
selected (for some examples where \( c \) is ‘other’ see [22, 23]). The following is a summary of the different simulated scenarios:

1. The first scenario simulated describes a situation where the prepared action has satisfactory predicted effects and therefore is executed; in this case both prior and retrospective awareness states occur. This scenario will be considered as the base case for the interplay between conscious and unconscious processes.

2. The second scenario simulated describes a situation where the prepared action has satisfactory predicted effects and therefore is executed but the awareness is absent (in other words merely an unconsciousness action). The strength of the action execution is lower as compared to the first scenario.

3. The third scenario simulated describes a situation where the prepared action lacks satisfactory predicted effects, and is therefore not executed: a no-go decision, or vetoing in unconscious form. Furthermore, the awareness state is almost absent due to the almost absent feeling.

4. The fourth scenario simulated describes a situation where a poor action prediction capability is modelled: the action effect is falsely predicted as satisfactory. This leads to a prior ownership state, which is sufficient to actually execute the prepared action. In this case a low retrospective ownership state and almost absent retrospective awareness state will occur, as the sensory representation of the effect stays low. This simulation is used to explain the basic cognition and behaviour of a schizophrenic patient.

5. The fifth scenario simulated describes a situation where two prepared actions exist for two input stimuli \((s_1, s_2, c_1, c_2)\) but one is relatively less positive compared to the other on predicted effects (difference is 0.2 in terms of the weight \( \omega_{10} \)). The one which is less positive (2\textsuperscript{nd} option) gets diluted over time in terms of \( \text{PA}(a_2), \text{SR}(b_2), \text{F}(b_2), \text{PO}(a_2, b_2, c_2, s_2), \) and \( \text{PAwr}(a_2, b_2, c_2, s_2) \) while the other prepared action gets executed and develops the retrospective awareness too.

6. The sixth scenario simulated describes a situation exactly as in the fifth scenario but in this case once the action with the strongest predictive effect (i.e. \( \text{SR}(b_1) \)) is executed it does not suppress the inputs \( s_2 \) and \( c_2 \). Therefore, once \( \text{EA}(a_1) \) is executed because of the existence of \( s_2 \) and \( c_2 \) the agent is preparing for \( \text{EA}(a_2) \) and successfully this will be performed. This confirms the agent capacity of cognitive control combining both conscious and unconscious processes.
7. The seventh scenario simulated describes a situation where the agent is prepared for an action by expecting a particular effect \( b_1 \), though it is actual effect after execution is different: \( b_2 \) (mismatch between the predicted and actual). As claimed by Haggard this scenario contributes to the idea that awareness of an action is a dynamic combination of both predictive and inferential sense-making. Having a strong predicted effect agent develops a strong prior awareness, but not sensing the same effect it leads to a poor retrospective awareness of that predicted effect. This phenomenon is important for the agent’s learning process through error correction.

8. The eighth scenario simulated describes a situation that can be considered as an early stage of a cognitive impairment for depressive symptoms. In this case the agent is preparing for two action options where one is having a positive feeling while the other is with a negative feeling (i.e. \( F(b_1) \) and \( F(b_2) \)). According to the biased nature on negative feelings the agent consciously affects selection of the negative one, though both options are identical from an action selection perspective (i.e., by having exactly the same values for all the parameters for each option except for \( \omega_{29} \)). Due to this conscious biased influence agent will execute the action with a negative feeling and it is considered to be that repeatedly performing these type of thoughts/tasks it will develop a negative mood that leads to a depression situation (in a long run).

Although these simulations will be presented in more abstract manner without relating them to real examples all the time, in general from a real world perspective the following example explanation of a scenario can be kept in mind:

- **Stimulus** \( s \) is that you need to withdraw some money.
- **Context** \( c \) is that you are in a bank and in front of an ATM.
- **Action** \( a \) is clicking on specific buttons to withdraw € 55.
- **Effect** \( b \) of action \( a \) is that you get € 55 cash.

This simple happy example scenario may have many variations such as:

- the person may doubt whether to directly withdraw € 55 or to first check the current balance in the account
- once the person has entered € 55 to withdraw, the system may inform the person that it does not have € 5 notes and ask to change the amount
- may be surprisingly ATM will return the card but without the money

Furthermore, when considering complex systems like Air Traffic Management there will be many examples that this model can be utilised (see [49] for more examples).
2.4.1. Selecting values for parameters of the model

Selecting parameter values for a dynamic cognitive model which consists with set of differential equations is a nontrivial research challenge. This will be even more difficult when there are no detailed numerical empirical data to use in this process, but only some characteristics of behaviours available in more fuzzy form (cf. [50]). Furthermore, another problem with computational cognitive models is that for different types of persons with different behaviours it may seem necessary to find a unique person-dependent set of parameters from scratch. As an alternative, if it is possible to identify a particular parameter value set which is able to demonstrate a variety of situations using only very minimum number of variations this would make the issue easier to handle. Therefore, in this work the focus has been on finding such a generic parameter value set that provides more confidence from the model validation perspective and its practical usage in future complex applications. The current model consists of many parameters: 39 weight values (for one option: \( k = i = l \)): \( \omega_i \), a time step size: \( \Delta t \), slow and fast speed factors: \( \gamma \), 17 steepness: \( \sigma_i \), and 17 threshold: \( \tau_i \). Table 4 presents connection weight values and Table 5 presents threshold (\( \tau \)) and steepness (\( \sigma \)) values used in configurations of simulations on this cognitive model. From these it is clear that the weight value set is generic and just changing very few weights (in most of the cases either one or two) have obtained the different expected behaviours.

The main challenge in this approach is that there is no real detailed data value set that can be compared to the output of the agent model to estimate parameters. Only certain features of the behaviour of each cognitive state are known for different scenarios, based on neurological and behavioural evidence from the literature (for example prior awareness should occur before the action execution and after prior ownership, there should be a dip in the sensory representation in-between predictive representation and inferential representation, et cetera). To identify the parameter values a systematic approach is used. For this approach it is a necessary condition to select multiple scenarios (minimum is 3 but having more will improve the quality of the results) which are interrelated from a functional point of view. For example, the first scenario is considered to be as a reference scenario and the second scenario is different from that just by achieving that awareness states are not developing (i.e. \( \omega_{24} = \omega_{25} = \omega_{34} = \omega_{35} = \omega_{36} = 0 \)). Furthermore, the third scenario handles a case in which a prepared action lacks satisfactory predictive effects (i.e. \( \omega_{24} = 0.2 \)), and the fourth scenario addresses poor action prediction capability (i.e. \( \omega_{18} = 0.2 \)); and so on (see Table 4). This interrelation among scenarios is very important for a minimum number of parameter changes enabling to identify a generic parameter value set for the model.
Table 4: Connection weight values used for cognitive agent model. In here if a value of a particular weight is empty for a scenario that means it is equal to the value of that in the Scenario 1. Furthermore if a value is ‘–’ then such a link was not existed for that scenario and furthermore ‘*’ presents that the particular link suppresses both its mapping inputs and all the remaining.

<table>
<thead>
<tr>
<th>Weights</th>
<th>Sce. 1</th>
<th>Sce. 2</th>
<th>Sce. 3</th>
<th>Sce. 4</th>
<th>Scenario 5</th>
<th>Scenario 6</th>
<th>Scenario 7</th>
<th>Scenario 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>ω₁</td>
<td>0.5</td>
<td></td>
<td></td>
<td>0.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ω₂</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ω₃</td>
<td>0.7</td>
<td></td>
<td></td>
<td>0.7</td>
<td>0.0</td>
<td>0.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ω₄</td>
<td>0.7</td>
<td></td>
<td></td>
<td>0.7</td>
<td>0.0</td>
<td>-</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>ω₅–ω₆</td>
<td>1.0</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₇</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₈–ω₉</td>
<td>1.0</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₁₀</td>
<td>0.8</td>
<td>0.2</td>
<td>0.6</td>
<td>0.6</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₁₁</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₁₂</td>
<td>0.5</td>
<td></td>
<td></td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₁₃</td>
<td>0.9</td>
<td></td>
<td></td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₁₄–ω₁₅</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₁₆</td>
<td>0.7</td>
<td></td>
<td></td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₁₇</td>
<td>0.0</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.1</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>ω₁₈</td>
<td>0.8</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₁₉</td>
<td>0.3</td>
<td></td>
<td></td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₂₀–ω₂₂</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₂₃</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₂₄</td>
<td>0.9</td>
<td>0.0</td>
<td></td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₂₅</td>
<td>0.8</td>
<td>0.0</td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₂₆</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₂₇–ω₂₈</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₂₉</td>
<td>0.5</td>
<td></td>
<td></td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₃₀</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₃₁</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₃₂</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₃₃</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₃₄</td>
<td>0.7</td>
<td>0.0</td>
<td></td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₃₅</td>
<td>0.7</td>
<td>0.0</td>
<td></td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₃₆</td>
<td>0.7</td>
<td>0.0</td>
<td></td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₃₇</td>
<td>0.7</td>
<td></td>
<td></td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₃₈–ω₃₉</td>
<td>0.8</td>
<td></td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In this parameter estimation approach the idea is as follows:
- First scenario is addressed and parameter values are calibrated to simulate its behaviour as identified through the literature
- Then by using the obtained parameter value set, by changing just a few (scenario-related) weight values it is checked whether the model with these parameter settings is able to generate the behaviour for the second scenario.
If this provides a simulation with a pattern as expected (without changes to the previously obtained parameter values, except for the changes particular to the current scenario) then it provides a good confidence on the current identified parameter value set,

But if not, then it is necessary to change the parameter values of the first simulation (based on the sensitivity of certain parameters on the required final output) until the behaviours for both simulations are satisfactory.

This approach is incrementally extended to each scenario until a generic parameter value set for all the scenarios has been identified. For any new scenario if any changes to the previously obtained parameter values are required, then all previously addressed scenarios are readdressed.

In the first few iterations it may challenging to identify a parameter value set, but over time the convergence is really fast and it will be experienced that the identified parameter value set is more generic and it is facilitating the necessary behaviours for simulations even without any changes to the obtained parameter value set.

In addition to the parameter values in Table 4 and 5 for the step size ($\Delta t$), slow speed factor ($\gamma$), and fast speed factor ($\gamma$) parameter values 0.25, 0.6, and 0.7 were used respectively for all the scenarios.

### 2.4.2. Scenario 1: Normal execution with ownership and awareness

The first scenario considered describes a situation where the context $c$ is the agent itself, and a stimulus $s$ occurs. The action effect $b$ of $a$ is considered positive for the agent and the awareness of action formation and execution will occur, together with generated prior and retrospective ownership states. The following execution trace will be expected from the agent here:

- External stimulus $s$ and context $c$ will occur and trigger preparation of action $a$.
- Based on the preparation state for $a$ the sensory representation of the (positive) predicted effect $b$ of $a$ is generated.
- Based on this positive predicted effect and the other states a prior ownership state for action $a$ is generated.

| Table 5: Threshold ($\tau$) and Steepness ($\sigma$) values used in configurations of simulations |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| $\sigma$ | $\tau$  | $\sigma$ | $\tau$  | $\sigma$ | $\tau$  | $\sigma$ | $\tau$  | $\sigma$ |
| 2.00    | 0.50   | 2.00    | 0.50   | 2.00    | 0.50   | 2.00    | 0.50   | 2.00    |
| 3.50    | 0.01   | 5.00    | 0.01   | 5.00    | 0.01   | 5.00    | 0.01   | 5.00    |
| 6.00    | 0.01   | 6.00    | 0.01   | 6.00    | 0.01   | 6.00    | 0.01   | 6.00    |
| PA      | F      | PO      | PAwr   | EA      | RO     | RAwr   | EO     |
| $(a_i)$ | $(b_i)$ | $(Y_{i,k})$ | $(a_i)$ | $(Y_{i,k})$ | $(b_i)$ | $(Y_{i,k})$ | $(b_i)$ | $(Y_{i,k})$ |
Prior ownership for action \( a \) is followed by the prior awareness; this is generated just before the action execution.

The prior ownership and prior awareness states for action \( a \) lead to actual execution of action \( a \).

The execution of \( a \) affects \( b \) in a positive manner and, via sensing the sensory representation of \( b \) and the feeling of \( b \).

At the same time the sensory representation of \( b \) is suppressed due to the prior self ownership state.

Based on the generated states, after the execution of action \( a \) the agent develops a retrospective ownership state.

Retrospective ownership of sensed effect \( b \) of action \( a \) is followed by retrospective awareness of action \( a \) and its effect.

Finally the agent communicates this ownership and awareness of it.

The simulation result of this scenario is shown in Figure 2. In this figure it is shown that (after sensing the stimulus) the agent triggers preparation of action \( a \) from time point 3 on (with a peak value of 0.74 around time point 55). Based on that the sensory representation of the predicted effect \( b \) of \( a \) is generated (through the as-if body loop with peak value 0.15 around time point 15 and through the body loop with peak value 0.59 after action execution around time point 50). This is followed by the feeling of \( b \) (through the as-if body loop with the peak value 0.19 and through the body loop with the peak value 0.62). These states contribute to the generation of a prior ownership state which starts to trigger at time point 5 and reaches a peak value of 0.77 around time point 57. After activating prior ownership,
prior awareness is developing, mainly upon the formation process of effect prediction $b$ of $a$, and its associated feeling. The prior awareness has started to pop up around time point 13 and has obtained peak value 0.71 around time point 45. As a result of the prior ownership and awareness states, the agent initiates the actual execution of action $a$ which propagates its effects through the (external) body loop. This clearly shows that prior awareness is just before the action execution (c.f. [4, 5, 8]) as the action execution process started at time point 15, and has its peak around time point 55, with maximum strength 0.92. Furthermore, it shows that via the body loop and the sensor state, the execution of action $a$ also affects the sensory representation of $b$ and the feeling of $b$. Therefore, the sensory representation $b$ of $a$ behaves as expected, by adding the sensed actual effect to the predicted effect, and the same effect is propagated to the feeling of $b$ too (c.f. [20, 40, 41]). Due to the action execution the agent develops a retrospective ownership state (starts at time point 26 with peak value 0.82 around time point 58), which is followed by a retrospective awareness state (starts at time point 29 with peak value 0.78 around time point 62). Finally, the agent communicates ownership and awareness of it for the performed action, based on the retrospective awareness and ownership states (with the maximum strength of 0.82 around time point 67). Note that when the stimulus is taken away (as explained in Section 3 this has been performed through an external suppressive mechanism through the orange colour arrows in Figure 1), all activation levels will come down to 0; they will come up again when a new stimulus occurs. Note that the numerical information related to the time scale or the peak values have not been coupled with actual brain signals but is only used as a frame of reference.

2.4.3. Scenario 2: Normal execution with ownership but without awareness

There are many situations in which human action formation occurs merely unconsciously, especially when related to habitual tasks [33]. As in the first scenario above also in this case the agent will experience that the prepared action has satisfactory predicted effects. Nevertheless, the agent will not develop any awareness state of the experienced feeling. The following execution trace will be expected from the agent for this scenario.

- External stimulus $s$ and context $c$ will occur and trigger preparation of action $a$.
- Based on the preparation state for $a$ the sensory representation of a (positive) predicted effect $b$ of $a$ is generated.
- Based on this positive predicted effect and the other states a prior ownership state for action $a$ is generated, but no prior awareness
- The prior ownership state for action $a$ leads to actual execution of action $a$.
- The execution of $a$ affects (via sensing) the sensory representation of $b$ and the feeling of $b$ in a positive manner
- At the same time the sensory representation of $b$ is suppressed due to the prior ownership state.
- Based on the generated states, after the execution of action $a$, the agent develops a retrospective ownership state, but no retrospective awareness.
- The agent does not communicate this ownership.

The simulation result for this process is presented in Figure 3. The agent starts to prepare for action $a$ at time point 2 and for this preparation a peak value of 0.52 is obtained. Together with the action preparation agent develops the sensory representation of predicted effect $b$ of $a$ (with peak value 0.15 based on the as-if body loop, and peak value 0.61 through the body loop) and the associated feeling of $b$ (with peak value 0.68). Based on this predictive information, the agent develops prior ownership from time point 6 on and with peak value 0.7. More importantly, in this simulation prior awareness has not developed. The developed states lead to performing the action $a$ which starts at the time point 14 and obtains peak value 0.43. The execution positively affects (via the sensor state) the sensory representation of $b$ and the feeling of $b$ (adding the sensed actual effect to the predicted effect). With this action execution effect the agent develops retrospective ownership with peak value 0.44. Finally, the agent does not communicate the ownership about the performed action (has a very low strength due to lack of awareness).

This cognitive behaviour trace is in line with the expectations and mainly when comparing it to Scenario 1 it demonstrates the possible impact of the awareness states. For example, when a prior awareness state occurs, it facilitates an enhancing

![Diagram](image-url)  
**Fig. 3:** Scenario 2: Executing an action with ownership and no awareness. In here 'Y_1' represents '$a_1, b_1, c_1, s_1$'.
effect on action preparation, sensory representation and feeling and also provides much smoother effects (as seen in Figure 2). Nevertheless, in this Scenario 2 these states have lower activation levels (also the action preparation state), that may be attributed to the absence of enhancing effects through prior awareness. Also this scenario highlights the strength of action execution with awareness: the action has executed with a relatively high peak value (0.92 vs 0.43). This might be explained as an influence from prior awareness to action execution. Nevertheless, according to the literature further research is required to conclude this; cf. [5]. From the computational perspective this at least confirms the model’s capability of action formation both with awareness and without awareness.

2.4.4. Scenario 3: Prepared action lacks satisfactory predicted effects

Humans are not always responsive to all the environmental stimuli (even in unconscious form). Nevertheless, there should be an explanation from the perspective of internal processes when receiving a stimulus why that stimulus does not lead to an actual action execution. This simulation provides the behaviour for such situation and explains how the lack of (positive) predicted effects generated through the as-if body loop relates to this. The following execution trace will be expected for this scenario:

- External stimulus \( s \) and context \( c \) occur and trigger preparation of action \( a \).
- Based on the preparation state for \( a \) only a weak sensory representation of predicted effect \( b \) of \( a \) is generated.
- Poor predictive effects will be reflected through the feeling state with very low activation.
- Based on this poor predicted effect \( b \) and the other states a low prior ownership state for action \( a \) is generated.
- Due to the poor prior ownership and low predictive feeling states, the agent does not develop adequate prior awareness.
- The low prior ownership state for \( a \) does not lead to actual execution of action \( a \); the action \( a \) can be considered vetoed
- The agent develops no retrospective ownership state for \( a \) and no retrospective awareness.
- The agent does not communicate ownership or awareness.

The simulation of this scenario is shown in Figure 4. The predicted effect is very low compared to the Scenario 1. This clearly shows that the action \( a \) triggered by stimulus \( s \) (which has an effect \( b \)) is not positive for the agent (in other words it is more like neutral to the agent in terms of the feeling): it leads to not getting any feelings out of it. Nevertheless, the prediction capabilities are assumed correct in this case, so no high level of \( b \) is correctly predicted for \( a \). As a result of this low prediction, the prior ownership state also stays at a very low level. Due to this, prior
Fig. 4: Scenario 3: Prepared action lacks satisfactory predicted effects. In here ‘Y_1’ represents ‘a_i, b_i, c_i, s_i’.

Awareness is not developed (stays in a very low level), which would be needed to strengthen the action execution. Therefore execution of the action also stays very low (below 0.1) and due to that, there is no retrospective ownership state, nor communication of ownership. Having a single weight value change (ω_{10}: PA(a_i) to SR(b_i) from 0.8 to 0.2) to obtain this behaviour from the reference Scenario 1 shows the coherent nature of action formation and higher order coupling as a process. This shows, from a complex systems simulation perspective, how the same agent model by limited variations in assigned parameter values demonstrates qualitatively different results. Furthermore, this confirms that the model has adequately adapted Damasio’s hypothesis: in the agent’s decision making process it has to assess the incentive value of the choices through an internal simulation process.

2.4.5. Scenario 4: Poor feelings of action prediction effects of a schizophrenic patient

In the previous scenario it was presented how the lack of satisfactory predicted effects will lead to a No-Go decision on action execution. In the current simulation scenario the focus is on poor feelings of predicted effects (the satisfaction level of the predicted effect is high but it does not adequately feel to the agent). In this situation the action effect b for action a_i, in principle is positive for the agent, like in the first scenario. Nevertheless the agent will not properly feel the effects of prediction. This is what assumed to happen in (at least some) patients with Schizophrenia [31, 40, 41]. Much evidence exists that relates this to poor emotional aspects in expression, experience and perception (mainly due to abnormalities in the
workings of Amygdala) [51, 52]. Schizophrenic patients often have the impression that their own actions are being created, not by themselves, but by someone from the outside [31]. The following execution trace will be expected for such a phenomenon:

- External stimulus $s$ and context $c$ occur and trigger preparation of action $a$.
- Based on the preparation state for $a$ the sensory representation of predicted effect $b$ of $a$ is generated.
- Lack of feeling of predicted effects will be experienced.
- Based on this predicted effects and its poor feeling a relatively low level of a prior ownership state for action $a$ is generated.
- Based on this low level of prior ownership and poor feelings, a relatively low level of prior awareness will be developed.
- This prior ownership and awareness level for action $a$ is still sufficient to lead to actual execution of action $a$.
- The execution of $a$ affects $b$ in a positive manner and (via sensing) the sensory representation of $b$ but still the felt feelings are weaker.
- Due to poor feelings agent will not develop adequate level of retrospective ownership and retrospective awareness.
- The agent does not communicate ownership or awareness for action $a$.

The simulation of this scenario is shown in Figure 5. In this case the agent has not fully felt the predicted effects of action $a$. After sensing the stimulus agent has triggered preparation of action $a$ at time point 3 (with a peak value of 0.52). Based on that the sensory representation of predicted effect $b$ of $a$ is generated (through the as-if body loop with peak value 0.25 and through the body loop with peak value 0.59) and followed by the feeling of $b$ (through the as-if body loop with peak value 0.08 and through the body loop with peak value 0.25). This clearly shows that the predicted effect has not been properly felt by the agent due to very low values for feeling state $F(b)$. Next these states contribute to generate a prior ownership which starts to trigger at the time point 6 with peak value 0.52. Together with the prior ownership the agent has experienced prior awareness with peak value 0.47 (this strength is relatively less compared to the same in the first scenario: 0.71). The prior ownership and awareness levels are much better in this case compared to the situation in the third scenario. Therefore, in contrast to the third scenario, these levels turn out high enough for the execution of the action. The maximum strength of the actual execution of action $a$ is 0.57 and this execution has positive effects which are sensed. Therefore, the sensory representation $b$ of $a$ behaves as expected after adding the sensed actual effect to the predicted effect. Nevertheless, the agent has again not properly felt the effects: it has not developed a sufficient strength for the feeling state. Due to these poor perceived effects of feeling the agent has not
developed an adequate level of retrospective ownership and no retrospective awareness. This behaviour can be interpreted as having some strength for the prior ownership and awareness for the action but no retrospective values for it. The agent may reach an internal conflict situation where the action seems not being created by itself, but by someone else. From the parameter values’ perspective this result has been achieved only with a single change from the first scenario on $\omega_{18}$ (from $\text{SR}(b_i)$ to $\text{F}(b_i)$) from 0.8 to 0.2. Nevertheless, the cognitive impairment behind a schizophrenic patient is more complicated than modelled here; for example, the impact of some other states (perception, attention, emotions, etc.) has to be considered as well. Therefore, it is required to extend the current model to provide more realistic cognitive behaviour for Schizophrenia. Nevertheless, the current behaviour already demonstrates some of the basics.

2.4.6. Scenario 5: Cognitive controlling when multiple action options compete

The fifth simulation scenario explains a situation when two action options exist (for the simulation two input stimuli were considered for this) and both have the potential to execute an action. The competition among the two options and the cognitive control through unconscious and conscious processes on this action selection is captured, resulting in execution of only one of the actions. For this behaviour two independent input tuples $(s_k, c_k)$ were used, occurring in parallel (nevertheless it is possible to use a single input that triggers two action options, but due to the requirements in the next scenario the mentioned approach was used). For
Fig. 6: Scenario 5: Cognitive controlling when multiple action options compete. In here ‘Y_1’ represents ‘a_1, b_1, c_1, s_1’ whereas ‘Y_2’ represents ‘a_2, b_2, c_2, s_2’.

Each option the same configurations were used as in the first scenario, except for few weights (see Table 4). In this scenario for the connection strength related with option 2 (i.e., k = i = 2) from the action preparation $a_2$ to its predicted effect $b_2$, a moderately low value has been selected: $\omega_{10} = 0.6$. Values for the other parameters were again the same as in Scenario 1 (cf. Table 4). More importantly, if in this situation at a given time only one input tuple occurs (either $s_1, c_1$ or $s_2, c_2$) then each action will execute according to the same behaviour as in the first scenario. The simulation of this scenario is shown in Figure 6. The following execution trace is expected from the agent in this case:

- External stimuli $s_1, s_2$ and contexts $c_1, c_2$ will occur and trigger preparation of actions $a_1$ and $a_2$ separately.
- Based on the preparation state for $a_1$ and $a_2$ the sensory representation of predicted effect $b_1$ of $a_1$ and $b_2$ of $a_2$ are also generated.
- Nevertheless, the agent will show strong effects on option $a_1$ and the rate of activation for preparation of $a_2$ will quickly slow down and disappear subsequently, due to suppression by the preparation for the other option.
- Based on this positive predicted effect and the other states for $a_1$, a prior ownership state for action $a_1$ is generated.
- Prior ownership for action $a_1$ is followed by the prior awareness, generated just before the action execution.
- The agent will develop neither prior ownership nor awareness for option $a_2$.
- Prior ownership and prior awareness states for action $a_1$ lead to actual execution of action $a_1$. 
The execution of $a_1$ affects $b_1$ in a positive manner and, via sensing, the sensory representation of $b_1$ and the feeling of $b_1$.

At the same time the sensory representation of $b_1$ is suppressed due to the prior self ownership state of it.

Based on the generated states, after the execution of action $a_1$ the agent develops a retrospective ownership state for action $a_1$.

Retrospective ownership of action $a_1$ is followed by retrospective awareness of action $a_1$.

Finally the agent communicates this ownership and awareness related with option $a_1$.

The behaviour captured in the Figure 6 is in line with the expected trace. Both action preparations (i.e. $PA(a_1)$ and $PA(a_2)$) are activated at time point 2 and, more importantly, with the same rate of activation strength until time point 15 (this highlights the non-biased effects in early action formation). Nevertheless, after time point 15 it is clear that option $a_1$ has maintained somewhat the same rate of activation for action preparation while for option $a_2$ the preparation activation strength has started to decrease, due to suppression by the preparation for option $a_1$, via the suppressive link with strength $\omega_{17}$. In parallel to action preparation the sensory representations and feelings for each option are also activated, but due to the assigned slightly lower value for weight link $\omega_{10}$ on option 2 the sensory representations and feelings for predicted effect of option 2 are not maintained, in addition to the effect through the suppressive link $\omega_{17}$ On the preparation state for $a_2$. Each option for action preparation independently suppresses its complements’, proportional to the current strength of each preparation (as in [42] for lateral inhibition). As $\omega_{10}$ for option 2 is slightly weaker, this contributes to the relatively a low value of $PA(a_2)$ in comparison with $PA(a_1)$. Therefore, through these unconscious mechanisms the activation level of preparation state $PA(a_2)$ becomes lower. Due to this bias in the action formation process, none of the other states related to the option 2 are activated. In contrast, for option 1 all the remaining states are activated in the same order as in the first scenario. More importantly, not having strong dips for sensory representation and feeling states as in the second scenario, this further highlights the conscious influence for action formation when compared to the aspects highlighted for the second scenario. The agent has developed both prior and retrospective awareness states with acceptable strength and finally has communicated the ownership and awareness specific to option 1.

As presented in the third section, this model includes some suppressive external links for the purpose of the scenario; for this scenario, once the action related to option 1 was executed it has suppressed all of the inputs, even those related to option 2 (i.e., $s_2$, $c_2$) and therefore all the states values become zero at the end of the simulation.
2.4.7. Scenario 6: Cognitive control when multiple action options compete: an extension of the fifth scenario

In the fifth scenario once the action $a_1$ related to option 1 was executed it has suppressed (or stopped) the input stimuli related with both the options (i.e., $s_1$, $c_1$, $s_2$, and $c_2$). In the current scenario the same identical setup was used but execution of action $a_1$ does not stop the input stimuli related with the second option. Therefore, from a cognitive perspective as the action option 2 has been suppressed by action option 1, once the effects of action option 1 have been realised, and due to that its triggers have disappeared, execution of action option 2 can get a second chance in the action formation process. In the current scenario this will be examined. The following execution trace will be expected from the agent in this case:

- External stimuli $s_1$, $s_2$ and contexts $c_1$, $c_2$ will occur in parallel and trigger preparation of actions $a_1$ and $a_2$.
- Based on the preparation states for $a_1$ and $a_2$ the sensory representation of predicted effect $b_1$ of $a_1$ and $b_2$ of $a_2$ are generated.
- Nevertheless, the agent will show stronger effects on option $a_1$ and the rate of activation for $a_2$ will quickly slow down and disappear subsequently.
- Based on the positive predicted effect and the other states for $a_1$ a prior ownership state for action $a_1$ is generated.
- Prior ownership for action $a_1$ is followed by prior awareness; this is generated just before the action execution.
- The agent develops neither prior ownership nor awareness for option $a_2$.
- Prior ownership and prior awareness states for action $a_1$ lead to actual execution of action $a_1$.
- The execution of $a_1$ affects $b_1$ in a positive manner and, via sensing the sensory representation of $b_1$ and the feeling of $b_1$.
- At the same time the sensory representation of $b_1$ is suppressed due to the prior self ownership state of it.
- Based on the generated states, after the execution of action $a_1$ the agent develops a retrospective ownership state for action $a_1$.
- Retrospective ownership for action $a_1$ is followed by retrospective awareness of action $a_1$.
- Finally the agent communicates this ownership and awareness related with option $a_1$.
- With the execution of action $a_1$ the agent suppresses the inputs $s_1$ and $c_2$.
- Due to the absence of inputs $s_1$ and $c_1$, preparation for action $a_1$ is not triggered anymore.
- The still existing inputs $s_2$ and $c_2$ still trigger preparation of action $a_2$. This is not suppressed by preparation of action $a_1$, since this is not activated anymore.
Based on the preparation state for action $a_2$ the sensory representation of predicted effect $b_2$ of $a_2$ is generated.

Based on this positive predicted effect and the other states a prior ownership state for action $a_2$ is generated.

Prior ownership for action $a_2$ is followed by prior awareness for action $a_2$, which is generated just before the execution of action $a_2$.

Prior ownership and prior awareness states for action $a_2$ lead to actual execution of action $a_2$.

The execution of $a_2$ affects $b_2$ in a positive manner and, via sensing the sensory representation of $b_2$ and the feeling of $b_2$.

At the same time the sensory representation of $b_2$ is suppressed due to the prior self ownership state.

Based on the generated states, after the execution of action $a_2$ the agent develops a retrospective ownership state for action $a_2$.

Retrospective ownership of action $a_2$ is followed by retrospective awareness of action $a_2$.

Finally the agent communicates this new ownership and awareness too.

The simulation of this scenario is shown in Figure 7. In this figure from time point 0 to (roughly) 80 the behaviour is exactly the same as in the fifth scenario, but around time point 80 due to the suppressive effect on inputs $s_1$ and $c_1$ the agent shows the effects of losing the suppression of action option $a_2$. Therefore the agent has again strengthens the preparation of action $a_2$ and subsequently the sensory representation of predicted effect $b_2$ of $a_2$, and the feeling of $b_2$. Prior ownership co-occurs with the above states (mainly due to its pre obtained activation strength). Followed by the prior ownership state the prior awareness state develops as expected. As a result of the prior ownership and awareness states, the agent initiates the actual execution of action $a_2$, which propagates its effects through the external body loop. This shows a possibility to get action $a_2$ executed, in contrast to Scenario 6 above where due to the action competition it is suppressed. The peak value obtained for the action execution is 0.88 (the same for option $a_1$ is 0.89); this clearly shows that both options have the same power in getting executed. The execution of action $a_2$ (via the body loop) has further effects too. Due to the action execution the agent develops the retrospective ownership state for action $a_2$, which is followed by a retrospective awareness state. Finally, the agent communicates ownership and awareness about the just performed action. This simulation shows the ability of action selection through competition mainly from an unconscious perspective. Having a slightly different weight value for $\omega_{10}$ (from PAwr($a_i, b, c, s_k$) to PA($a_j$)) on action option 2 the same process can be simulated to demonstrate the effects of conscious cognitive control as a top-down effect. Furthermore, it is also possible to combine both of these effects in a simulation.
2.4.8. Scenario 7: Mismatch between the predicted and actual effects of an action

In most of the previous scenarios the predicted effect always has positively affected action execution. Nevertheless, it may not always be the case and as pointed out in Section 1 (with the example of learning how to ride a bicycle) there may be a difference between what is predicted and what actually occurs. In this scenario this phenomena will be simulated. For this simulation a single input is considered but which triggers preparation of an action $a_1$ for which two options for an effect are considered: the first is $b_1$ which is the predicted effect of $a_1$, whereas the second $b_2$ is not predicted. The second option indicates what actually will occur after execution of $a_1$ (see Table 4). In Table 4 the weight changes for this case have been highlighted for each effect option separately. For the effect option 1 weight $\omega_3$ (i.e., from EA($a_1$) to WS($b_1$)) was set ‘0’ as the effect of actual execution is not $b_1$ and the same weight for effect option 2 was set ‘0.5’ to facilitate the different non-predicted effect of the action (the same explanation for $\omega_4$). The weight from WS($b_2$) to SS($b_2$) (i.e., $\omega_7$ for option 2) was set 0.5 to facilitate the effects of actual sensing (this value can be further increased but 0.5 was selected merely to highlight an average effect in the sensing). Also according to the formation of this scenario at the beginning the agent will only predict the effect $b_1$ of action $a_1$, but not $b_2$. Therefore, the weight $\omega_{10}$ (i.e. PA($a_1$) to SR($b_2$)) was set ‘0’ (the same explanation applies for weight $\omega_{12}$). The weight $\omega_{30}$ from EA($a_1$) to RO($a_1$, $b_1$, $c_1$, $s_1$) was set ‘0’.
as there will not be any retrospective ownership for action $a_1$. Furthermore, the weight $\omega_{32}$ for effect option 2 (i.e., from $PO(a_1, b_2, c_1, s_1)$ to $RO(a_1, b_2, c_1, s_1)$) was set ‘0’ as there is no prior ownership on action option $a_2$. Additionally, the link from $EA(a_1)$ to $RAwr(a_1, b_1, c_1, s_1)$ (i.e., $\omega_{34}$) was also set ‘0’. Also as there is no prior awareness related with effect option $b_2$, weight $\omega_{37}$ (i.e., from $PAwr(a_1, b_2, c_1, s_1)$ to $RAwr(a_1, b_2, c_1, s_1)$) was also set ‘0’. The following execution trace is expected from the agent:

- External stimulus $s_1$ and context $c_1$ occur and trigger preparation of action $a_1$.
- Based on the preparation state for $a_1$ the sensory representation of a (positive) predicted effect $b_1$ of $a_1$ is generated.
- Based on this positive predicted effect and the other states a prior ownership state for action $a_1$ is generated.
- Prior ownership for action $a_1$ is followed by prior awareness, which is generated just before the action execution.
- Prior ownership and prior awareness states for action $a_1$ lead to actual execution of action $a_1$.
- The execution of $a_1$ does actually not affect $b_1$ but a different effect $b_2$. Therefore, through sensing the agent will develop a sensory representation of $b_2$ and the feeling of $b_2$.
- Based on the generated states, after the execution of action $a_1$ the agent may develop a low retrospective ownership state for action $a_1$ with effect $b_2$ due to the conflict between predicted effect and sensed actual effect.
- Retrospective ownership of action $a_1$ with effect $b_2$ is followed by retrospective awareness of action $a_1$ with effect $b_2$.
- Finally the agent may not properly communicate this ownership and awareness (depend on the context).

The simulation of this scenario is shown in Figure 8. As expected the agent triggers preparation of action $a_1$ at time point 3 (with the peak value of 0.73). Based on that the sensory representation of predicted effect $b_1$ of $a_1$ is generated (through the as-if body loop with peak value 0.23), followed by the feeling of $b_1$ (through the as-if body loop with the peak value 0.30). Next these states contribute to generate a prior ownership (for $a_1$ with effect $b_1$) which starts to trigger at the time point 5 with peak value 0.84. After activating the prior ownership, prior awareness develops, mainly upon the formation process of effect prediction $b_1$ for $a_1$. The prior awareness starts to pop up around time point 12 and obtains peak value 0.85. As a result of the prior awareness and ownership states, the agent initiates the actual execution of action $a_1$ which propagates its actual effects through the (external) body loop. The action execution starts at time point 13 and its maximum strength is 0.97.
Fig. 8: Scenario 7: Mismatch between the predicted and actual effects of an action. In here ‘Y_1’ represents ‘a_1, b_1, c_1, s_1’ whereas ‘Y_2’ represents ‘a_2, b_2, c_2, s_2’.

In Figure 8 it is clearly shown that the execution of action a_1 does not affect b_1 via sensing. Instead it has a different actual effect b_2; through sensing it affects the sensory representation and the feeling of b_2. Therefore, the sensory representation b_1 of a_1 does not behave as expected by adding the sensed actual effect to the predicted effect and the same effect does not propagate to the feeling of b_1 (c.f. [20, 40, 41]). Two activations: that do occur are of the sensory representation of b_2 and feeling for b_2. They emerge just at that point in time as they were not there at the beginning. These new states provide satisfactory activation levels but contribute for relatively low retrospective ownership and awareness states (mainly due to the nonexistence of influences of prior ownership and awareness states, respectively). Subsequently due to these poor retrospective ownership and awareness states agent does not properly communicate the action and its effect. Poor communication was noted experimentally in [53] through a card game together with a covert exchange to facilitate the conflict between predicted and actual outcome. Therefore, also from experimental findings this poor communication can be justified. Furthermore, in simulation scenario 1 when there is no conflict between predicted and occurring effect, prior awareness on sensed effect has initiated at time point 35 whereas in this simulation it occurs around time point 50. Therefore, the temporal gap between the action and its perceived sensory outcome is less when the awareness is pre-existing but it is more when the (prior) awareness is not involved, which is observable from these simulation results; this is referred as intentional binding [15].
This behaviour has shown the effects when there is a conflict between what predicted vs actual. In future work this will be especially useful when learning should incorporate with the model in a dynamical form. Having this conflict between sufficiently large prior awareness and a very low retrospective awareness can be further explored with more neuroscientific evidences. This is useful for adaptive behaviours especially for situation awareness driven applications.

2.4.9. Scenario 8: Cognitive impairment for shifting to a positive feeling

In real life agents will experience both positive and negative situations. Nevertheless, if an agent always sticks to negative thoughts while suppressing the possible positive thoughts, in the long run this will lead to a depression. In this scenario the agent is preparing for two action options where one leads to a positive feeling $F(b_1)$ while the other leads to a negative feeling $F(b_2)$. According to the bias to negative feelings, the agent tends to select the negative action though both options share exactly the same weight values for all weights on each option, except for $\omega_{29}$ which is for prior awareness (which introduces a conscious intervention on action execution). For the weight $\omega_{29}$ on the first (positive) action option 0.2 was used and for the second (negative) option 0.8. Except this change all the other parameter values are identical and through this only the impact from biased conscious awareness towards a particular action is modelled. If the agent is executing the action which is associated to negative feeling (though at the same time a possible positive action also exists) this will affect the agent’s mood (the mood of the agent has not been in the scope of this model but will be a future work). Having a negative mood all the time together with this biased awareness towards such negative actions will take the agent to a cognitive disorder state called depression. A healthy agent will have the ability to shift in a relatively short time period again to positive actions by correcting the biased influence. Therefore this model can be used to simulate a depression situation also, but with further improvements in the model. For the biased cognitive impairment on negative actions through awareness the following trace is expected:

- The external stimulus $s_1$ and context $c_1$ will occur and trigger preparation of actions $a_1$ and $a_2$.
- Based on the preparation state for $a_1$ the sensory representation of predicted effect $b_1$ of $a_1$ is generated and in the same way for $a_2$ the sensory representation of predicted effect $b_2$ of $a_2$ is generated.
- It is assumed to be that both effects positively contribute to the action formation process while the effect $b_1$ has a positive associated feeling, and the effect $b_2$ has a negative associated feeling.
- Based on this positive predicted effect and the other states a prior ownership state for action $a_1$ and $a_2$ are independently generated.
– Prior ownership for action \(a_1\) is followed by the prior awareness of \(a_1\) and the prior ownership for action \(a_2\) is followed by the prior awareness of \(a_2\).

– Nevertheless the prior awareness of \(a_2\) will dominate the action formation process due to the biased competition.

– Due to this the states related to action \(a_1\) will lose their activation while the states related to action \(a_2\) will continue as the selected options.

– Prior ownership and prior awareness states for action \(a_2\) lead to actual execution of action \(a_2\).

– The execution of \(a_2\) affects \(b_2\) in a positive manner through the sensing.

– The agent develops sensory representation of \(b_2\) and the feeling of \(b_2\) in line with what is predicted.

– Based on the generated states the agent develops a retrospective ownership state for action \(a_2\) with effect \(b_2\).

– Retrospective ownership of action \(a_2\) with sensed effect \(b_2\) is followed by retrospective awareness of action \(a_2\) with effect \(b_2\).

– Finally the agent communicates this ownership and awareness.

The simulation of this scenario is shown in Figure 9; it is in line with the expected trace. The agent has initiated two action preparations once the input stimuli are received. Both \(PA(a_1)\) and \(PA(a_2)\) are activated at time point 3 and almost with the same activation speed. Parallel to this the agent also initiates the sensory representation and feelings for effects \(b_1\) and \(b_2\) for actions \(a_1\) and \(a_2\) respectively. Furthermore, the rates of activation for these four states are also almost the same at the very beginning (until prior awareness states pop up). Followed by these states prior ownership states are also activated for both options and then subsequently the prior awareness states for each option. Roughly at time point 10 the prior awareness states start to emerge and from that very moment the states related with the option 1 show decline effects. For example, the action preparation states initially have same speed but after the time point 10 state \(PA(a_1)\) is losing the speed whereas \(PA(a_2)\) maintains the same momentum and reaches maximum value 0.75. Similar to this \(SR(b_1)\) reaches peak value 0.11 through the as-if body loop (i.e., through the predictive process) whereas \(SR(b_2)\) reaches peak value 0.23 through the predictive process and 0.63 through the body loop. Furthermore, \(F(b_1)\) reaches peak value 0.14 through the as-if body loop whereas \(F(b_2)\) reaches peak value 0.30 through the predictive process and 0.66 through the body loop. Moreover, \(PO(a_1, b_1, c_1, s_1)\) reaches peak value 0.15 whereas \(PO(a_2, b_2, c_2, s_2)\) reaches peak value 0.68. Also \(PAwr(a_1, b_1, c_1, s_1)\) reaches only a 0.09 value but \(PAwr(a_2, b_2, c_2, s_2)\) reaches 0.68.

These results clearly show the impact from conscious intervention for the action formation process (all the other unconscious related suppressive parameter values are identical for both options and therefore no bias was introduced by unconscious
processes). Therefore, due to this bias the negatively associated action option $a_2$ has executed with peak value 0.92. Furthermore, it is shown that the execution of action $a_2$ (via the body loop) also affects in positive manner via sensing, the sensory representation of $b_2$ and the feeling of $b_2$. Due to these the agent develops $RO(a_2, b_2, c_2, s_2)$ at time point 30 with peak value 0.79 and $RAwr(a_2, b_2, c_2, s_2)$ at the time point 34 with peak value 0.76. The agent does not develop any retrospective effects associated to action option $a_1$. Finally, the agent has strongly communicated the ownership and awareness on action $a_2$.

2.5. Discussion and Future Work

Computational modelling is considered an important pillar for the development of cognitive science and related disciplines [54-56]. Moreover, the developments in brain imaging and recording techniques also strongly contribute to more and more focused phenomena explored in the cognitive, behavioural, affective, and physiological research areas. Nevertheless, there is room for a more generic, compound process to explain a wide range of cognitive functionalities by aggregating many of these local but highly influential information from the computational perspective [57]. For example, Ron Sun in [57] has stated that: “integrative computational cognitive modelling may serve in the future as an antidote to the increasing specialization of scientific research” [57], pp. 14. On the other hand, there are many implications of these hypotheses (or evidences) and it would be beneficial if there was a mechanism that can be used to scrutinize these
ideas or hypotheses by using this as a workbench in much more abstract and global level (cf. [55]). Additionally, the human brain and its phenomena are immeasurably complex systems/processes that involve many uncountable factors that make experiments not always coherent with reality. Nevertheless, having computational models enables to uplift the progress of understanding these processes in a broader level as a multidisciplinary approach (cf. [54]).

From the Artificial Intelligence perspective more and more complex systems related problems are addressed that include human cognitive aspects. For example, situations related to air traffic management, stock market analysis, business processes, human awareness of energy usage, cognitive impairments and various medical disorders, how a certain form of therapy can have its effect on a patient, etc. In many such situations it is not practical to create real experimental setups to analyse the emergence of problems. Nevertheless, the importance and significance of such analysis is essential from a safety, performance, and health perspective. Multi-agent based simulation approaches have been noted as the potential solution for this [58] by considering agent based simulations. Such simulations can capture emergent phenomena, providing a natural description of a system, and its flexibility: essentials for such complex system research. Even though agent-based simulations are a promising technology for this, there are many problems associated with it in terms of more realistic models for such agents to behave or perform. Often non nature inspired, simple heuristically determined rule-driven agent simulations are used for this, although reality is more complicated and far away from those simplifications. It needs more realistic representations and analysis in much closer to the natural situation (cf. [58]). Therefore, in agent systems more realistic computational models that make use of the latest neuro-cognitive findings can be used to simulate agent behaviour in a more realistic manner. Especially when it comes to the systems that include human cognition factors in dynamic systems, nature inspired cognitive computational models have more power to provide realistic results. Therefore, computational cognitive modelling as a multidisciplinary approach has many benefits for both cognitive science and artificial intelligence.

This paper presented a computational cognitive model for action awareness focusing on action preparation and performance by considering its cognitive effects and affects from both prior and retrospective perspective relative to the action execution. It is a fundamental research question what is the human cognition behind the action selection and how this is related to conscious and unconscious elements. Mainly for two hypotheses or claims attention has been obtained by the community, as presented in Sections 1 and 2. However, it is not yet clear what is the exact process behind this human cognition. The first hypothesis (by Benjamin Libet and others) claimed that humans may prepare for and perform actions without being
conscious of these preparation and execution processes, and the awareness of motor intention of this action is not causing the behaviour, but comes after the action preparation and relatively just before the action execution time [4–8]. The second hypothesis (by Haggard and co-workers) claims that awareness of an action is a dynamic combination of both prior awareness and retrospective awareness through predictive motor control and inferential sense-making relative to the action execution respectively [20, 21]. Furthermore, the intentional binding effect [15] shows the impact of awareness on action selection and this contributes to the second hypothesis to bring out the influence of awareness on action selection. Although these two hypotheses seems to contradict each other from the semantic point of view, by other research (cf. [5]) from a pragmatic point of view it seems hard to generally reject any of these claims on the basis of the available empirical evidences. This paper utilizes both ideas into a compound process (together with other supportive processes) and scrutinizes the behaviour through related scenarios. This work was not conducted from scratch, but adopts parts of the model presented in [23], mainly for the mechanisms of action ownership and other unconscious states/processes (e.g., action preparation, sensory representation, effect prediction process, mirroring, etc.). Having that previous model which was validated through simulations mainly for unconscious action formation, in this paper its scope is further extended to incorporate the conscious aspects related to action selection. The main research questions for this work are:

1. How does the internal prediction process shape or contribute to the (prior) awareness of the action?
2. How does the inferential sensemaking shape or contribute to the (retrospective) awareness of the action execution?
3. How does the awareness contribute to action execution?
4. What is the relation and interplay between conscious and unconscious action formation through action ownership and relevant awareness states?

Each research question is explored from the cognitive science perspective and analysed through the modelling perspective (with simulations) to isolate a working definition. For the internal prediction process the hypothesis of Damasio’s as-if body loop is used as a basis, as previously presented in [23]. In the current model, it is further extended by embedding an unconscious process referred as the lateral inhibition (see [42]) to strengthen the competition among action options through the as-if body loop. With that new addition the unconscious action prediction process is coupled with the prior awareness state (as a higher order cognitive state) to facilitate the conscious aspects as proposed by Haggard et al.: predictive processes are also playing a role in action awareness (e.g., see [20, 21]). Also given empirical
evidence to support the idea that awareness of motor intention for an action comes after the action preparation and relatively just before the action execution time (cf. [4–8]), in this model it is ensured that the awareness states are always higher order cognitive states which are not getting affected by most of the low level cognitive states. Predictive processes anticipate effects of each action option and lead to a competition to get selected a GO signal. The basic decision making is assumed to be based on a feeling-related valuations associated with the effects of each action option. These internally predicted feelings work as a bottom-up feedback to develop a coherent conscious experience of action selection, mainly on action options that have strong predictive effects, therefore, actions with poor predictive effects do not get any conscious attention. Therefore, prior awareness state is only affected by the feeling and prior ownership states. The simulation results have confirmed that when there is a prior awareness it always appears just before the action execution.

The same approach is applied to the second research question, also to isolate a working definition. Once the feedback sensory information on the effects of the actual action execution is available and is integrated with the predicted effects as suggested in [20, 40, 41] to evaluate the binding of what is predicted and the actual effect. Through this sense making process the agent will experience the retrospective effects as presented in previous work [23], mainly for acknowledging authorship of an action, reflection on one’s own functioning, personal learning and development. This was further extended with a new abstract cognitive state name called retrospective awareness which is responsible for more conscious interpretation of retrospective effects of the action execution. Having empirical evidences for the contribution of inferential processes on action awareness from Haggard and co-workers’ experiments (cf. [20, 21]) this extension can be justified from the cognitive neuroscience perspective. Furthermore, through simulations on relevant scenarios this extension was validated. Especially the seventh scenario clearly shows that when there is a mismatch between what is predicted and what actually occurs, the two steps sigmoid behaviour on the sensory representation and the feeling states is not observed, whereas a poor retrospective awareness on the effect was predicted, as highlighted in [20]. When considering the first scenario and the seventh, this may even contribute for the observation of intentional binding (cf. [15]) mainly through the feeling state: increasing the same feeling through inferential information and forming a completely new feeling through the inferential process. In the seventh simulation the agent does not aggregate the prior awareness to the retrospective awareness and therefore the agent will not be consciously able to communicate the conflict or mismatch. This also shows the subjective time effect found in the intentional binding: when there is no conflict between predicted and occurring effect, the prior awareness on sensed effect has initiated earlier than when a conflict occurs. Therefore, the feeling of the temporal gap between the action and
its perceived sensory outcome can be explained through the retrospective awareness [15]. To have these results in this model the retrospective awareness states were mainly affected by four states: action execution of \( a_i \), feeling of effect \( b_i \), prior awareness for action \( a_i \) with \( b_i, c_k, \) and \( s_k \), and prior ownership for action \( a_i \) with \( b_i, c_k, \) and \( s_k \).

The third research question shows many thoughts within the cognitive neuroscience community. It suggests that awareness is not a cause for an action execution but it seems like an after-effect of a set of unconscious cognitive processes leading to the action [1, 4–8]. The other idea shows that an additional influence through awareness may inject some bias or effect on decision making especially with the findings of intentional binding [15, 21, 30, 36]. There are empirical evidences to support the second claim (at least for some inputs), emphasising why it is hard to accept the first claim due to its non moderate statistical strength on empirical evidences: in general only \( \sim 60\% \) accuracy in experimental data (see [5]). Therefore, this model includes both features: awareness is not required for action execution and awareness may play a role in action execution, depending on the specific scenario. In other words, the model handles both the purely unconscious action formation and the hybrid form of action formation with both conscious and unconscious elements. Furthermore, through the simulation results also this was validated. Especially the first and second scenarios provide information to show that the agent is able to execute an action with and without awareness. The eighth scenario presents a situation in which biased awareness may transform a healthy person into a cognitively impaired position. By continuing the cognition behind the eighth scenario for a longer period of time this can even explain the effects of a patient suffering from a depression (this will be considered as a future work). Therefore, with different settings and scenarios the model is capable to demonstrate the contribution of awareness from zero to high. Therefore, the model facilitates a good spectrum to represent the effects of awareness on action selection. There are situations, for example like flight or fight situations, which mostly show the unconscious action formation side of the spectrum that includes very quick and strong action executions (cf. [59–61]). According to the second scenario it is clear that when agent is purely performing in an unconscious mode the strength of action execution is relatively weak. Therefore, it may be possible to add improvements to this model and its settings, especially on the unconscious perspective including the emotion related effects. Further information on the interplay between bottom-up and top-down processes seems to be useful for such further improvements of this model (see [59–62]).

The fourth research question is realised as an aggregation of the other three questions. The interplay between conscious and unconscious action formation is mainly realised through having effects from feeling and ownership states on the
awareness states (unconscious to conscious), and effects from prior awareness states to the action preparation and action execution states (conscious to unconscious). The first links (bottom-up) play a role to pass low level information to develop awareness of what is going on in a high level form, whereas the second type of links (top-down) contribute to inject some bias or excitement to intentionally drive the action formation. When there is poor activation of the bottom-up links agent is unable to develop prior awareness and therefore, as presented in the second scenario the agent can perform the action selection in purely unconscious form by the feeling of ownership. When the agent is having sufficient activations in bottom-up links, awareness develops and shows how that leads to an action execution, as in the first scenario. Also particularly in the eight scenario the role of top-down links demonstrates the power of an intentional focus on action selection. The third simulation highlights why agents are not always performing a task even in unconscious mode. This shows that even in unconscious form it is necessary to have a sufficiently large satisfactory predicted effect to have a GO signal. Therefore, this shows the mechanism of vetoing in unconscious form. The cognitive control is a useful process to explain the interplay between conscious and unconscious action formation. The fifth and sixth simulations present the role of cognitive control. In the fifth simulation it shows how a particular option is getting suppressed by the other option through cognitive control just by having a value of 0.2 difference for the links between action preparation to sensory representation for two options. The sixth simulation shows that once the dominant option of the fifth simulation completed, the suppressed action option emerges due to not having the effects of cognitive control and, more importantly, with strong activations for each state. Through this it is clear from the simulation perspective that the model can have different configurations to facilitate different behaviours both from the unconscious and the combination of conscious and unconscious forms. Also as mentioned in the other 3 research questions the cognitive neuroscience basis behind this model was inspired by the experiments and evidences found in the literature.

Having interesting simulation results for many scenarios still there is more to improve on this work. It was possible to isolate a generic parameter value set that worked for 8 simulations. A generic parameter value set provides more confidence from the model validation perspective and its practical usage in future complex applications. Nevertheless, it is not a trivial task to find such generic parameter value set. The approach used to identify this parameter value set was explained in Section 4.1, but it is not a fully automated process. Due to the complexity of the human brain and limitations in measuring techniques of human cognition there are no detailed numerical empirical data to use in this process, but only some characteristics of behaviours available in more fuzzy form. Due to this issue, it is not possible to directly use parameter estimation techniques available for dynamic
systems. Therefore, it is useful to explore the parameter estimation techniques particular to the characteristics of these types of work. Furthermore, from the cognitive neuroscience perspective, this model has many more areas to explore both in conscious and unconscious levels. In the third scenario, it shows the vetoing process in unconscious form, but the same process with awareness which is referred as intentional inhibition is future work for this model. Human awareness has its specializations, for example, emotional awareness, situation awareness. Working processes behind these concepts more complicated and need further research to incorporate those into this model.

Finally, this model may be useful in many applications. Especially for agent-based simulations on complex systems that need action selection related to cognitive aspects. Also, this model can be used as a basis for subsequent work in developing ambient agent systems able to monitor, analyse and support persons trying to develop a healthy lifestyle. If such systems have such a model of the underlying human processes, they can use this to have a deeper understanding of the human.

References


http://doi.org/10.1177/1073858407299288

http://doi.org/10.1162/089892999563607

http://doi.org/10.1080/13546800143000212


http://doi.org/10.1016/j.bica.2012.07.005

http://doi.org/10.1093/mind/os-IX.34.188

http://doi.org/10.1142/S0218213007003357


http://doi.org/10.1016/j.pneurobio.2005.11.005

http://doi.org/10.1177/0963721410377599

http://doi.org/10.1126/science.1111709

http://doi.org/10.1111/j.1756-8765.2010.01092.x

http://doi.org/10.1111/j.1756-8765.2008.01003.x

http://doi.org/10.1111/j.1756-8765.2012.01206.x

http://doi.org/10.1073/pnas.082080899

http://doi.org/10.1002/hbm.20037

http://doi.org/10.1093/scan/nsq103

http://doi.org/10.1111/j.1467-9280.2009.02459.x

http://doi.org/10.1016/S0028-3932(02)00045-3
Chapter 3
Modelling Intentional Inhibition of Actions\textsuperscript{1}

Dilhan J. Thilakarathne, Jan Treur

Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.
e-mail: d.j.thilakarathne@vu.nl, j.treur@vu.nl

Abstract: Inspired by cognitive and neurological literature on action ownership and action awareness, in this paper a computational cognitive model for intentional inhibition (i.e., the capacity to voluntarily suspend or inhibit an action) is introduced. The interplay between (positive) potential selection of an action, and (negative) predicted impact of this action is addressed. Neurological evidence indicates that this interplay of positive and negative influences on action selection has contributed to make a homo sapiens into a social being. In this process performative and constitutive desires are used to differentiate the influence for action preparation and intentional inhibition with the relevant supportive states: ownership and awareness states. Intentional inhibition is different from action inhibition processes related to an externally guided stop signal. In intentional inhibition process mainly an internally guided action is inhibited through an internally guided stop signal. The proposed model was validated through five scenarios by using a generic parameter value set which is identified through a systematic analytical approach. These scenarios cover the situations in which the agent is: 1) intentionally inhibited action with satisfactory predicted action effect, 2) unable to intentionally inhibit an action with satisfactory predicted action effect, 3) influence of impact prediction on intentional inhibition process, 4) action selection and inhibition in a fully unconscious situation, and 5) impact of the lack of constitutive desires in the effect prediction process. The introduced computational model provides a basis for application domains concerning decision making, behavioral management, emotional control, and simulations for clinical disorders and therapies for them.

Keywords: Intention inhibition, computational cognitive model, performative and constitutive desires

\textsuperscript{1} This chapter was published as:
http://doi.org/10.1016/j.bica.2015.07.001

Which is an extended work of the following conference paper:
http://doi.org/10.1016/j.sbspro.2013.10.205

(Received the Best Paper Award at the conference)

The names of the authors are ordered alphabetically reflecting the comparable contribution of each author.
3.1. Introduction

Desires or feelings may be assumed to govern preparation and execution of your actions. However, in certain situations spontaneously deciding not to do a prepared action may be more vital. The term intentional inhibition refers to the capacity to voluntarily suspend or inhibit an action; (Filevich, Kühn, & Haggard, 2012; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012; Kühn, Haggard, & Brass, 2009; Brass & Haggard, 2008, 2007; Haggard, 2008). An example of this (Filevich, Kühn, & Haggard, 2012) is that you are writing an email to your boss, and just before you click the ‘Send’ button, you seem to hear a voice in your head that says ‘do you really want to send that?’, and you hold back. Intentional inhibition may take place as part of the process of preparing for and deciding about actions to be performed. This capability is considered of importance for successful social interaction and personal development. In the explanation of intentional inhibition an important role is played by distinction between performative desires and constitutive desires; for example, a desire to take fast food may be a performative desire, whereas a desire to take healthy food may be a constitutive desire. Some actions may satisfy performative desires, and the same or different actions may satisfy constitutive desires. It is this difference where intentional inhibition comes in: if an action satisfies a performative desire but not a constitutive desire, intentional inhibition may prevent the action from becoming executed. Through this, persons can refrain from doing something they are tending to do.

Among the many networks in the human brain, a specific network addresses inhibition of intentional actions (Brass & Haggard, 2007, 2008). This network is mainly centered in the dorsal Fronto-Median Cortex (dFMC) and activates when a cognitive conflict occurs on whether you should prevent yourself from doing what you are about to do. There is recent evidence on the involvement of dFMC in the suppression of emotions or desires which may not directly lead to an immediate behavioral output (see Lynn, Muhle-Karbe, & Brass, 2014). Furthermore, this network contributes to generate conscious awareness on intending (Lau, Rogers, Haggard, & Passingham, 2004), which makes the intentional inhibition a more conscious process. The skill of intentional inhibition is lacking among gamblers; they have the impression that if they keep on gambling they can win back the money that they have lost. This phenomenon of repeated gambling to recover losses is called ‘loss-chasing’ (Campbell-Meiklejohn, Woolrich, Passingham, & Rogers, 2008). There is experimental evidence that shows how loss-chasing is correlated with impairments in brain networks related to intentional inhibition: ventromedial Prefrontal Cortex (vmPFC) and subgenual Anterior Cingulate Cortex (sgACC) activates in the chasing period whereas dorsal Anterior Cingulate Cortex (dACC) and Anterior Insula Cortex (AIC) activates when quitting the chase (see Campbell-
Meiklejohn, Woolrich, Passingham, & Rogers, 2008). Moreover, as dACC and AIC are overlapping with the dFMC, this further contributes to the idea that intentional inhibition emerges through a particular network and not enabling this neural system makes it harder to stop (or suppress the intending action) loss-chasing for gamblers (Brass & Haggard, 2008). Overall, individuals are assumed to have an internal capacity to inhibit actions, and having problems in that capacity may lead to many problems such as drug and alcohol abuse, excessive eating, smoking, or excessive gambling.

In this paper a computational cognitive model for intentional inhibition is introduced based on literature such as (Filevich, Kühn, & Haggard, 2012; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012; Kühn, Haggard, & Brass, 2009; Brass & Haggard, 2008, 2007). This computational model is inspired by cognitive and neurological literature about intentional inhibition, action ownership and action awareness. The notions of ownership and awareness of an action have received much attention in the recent cognitive and neurological literature (see Thilakarathne & Treur, 2015). This paper extends the work published in Thilakarathne and Treur (2013a) by incorporating more neuroscientific evidence and new simulation results. The introduced computational model provides a basis for application domains concerning decision making, behavioral management, emotional control, and simulations for clinical disorders and therapies for them.

3.2. Conceptual Basis

An important role both in the execution decisions for an action, and in its attribution, is played by the prediction and valuation of the (expected) effects of the action, based on internal simulation starting from the preparation of the action; e.g., (Wolpert, 1997; Haggard, 2008). If these predicted effects are valued as satisfactory, this may entail a ‘go’ decision for the execution of the action, thus exerting control over action execution. In contrast, less satisfactory predicted effects may lead to a ‘no go’ decision. A related element, put forward, for example, in Moore and Haggard (2008), is the notion of action awareness. In Thilakarathne & Treur (2013b, 2015) based on specific neuro-cognitive evidence a computational model for these processes is contributed in relation to concepts such as ownership, awareness, feeling, action preparation, sensory representation and communication. This conceptual basis covers processes behind intentional inhibition. Intentional inhibition is different from inhibition driven by external stimuli where external information motivates to stop the behaviour (see Diamond, 2013). Both action initiation and inhibition can be guided by externally and/or internally triggered processes, which provides four combinations (the provided examples are from (Schel, Scheres, & Crone, 2014)):
(1) externally guided action inhibited through externally guided stop signal: e.g., stop walking when a green traffic light suddenly turns red.

(2) externally guided action inhibited through internally guided stop signal: e.g., resisting the impulse to take another biscuit from the biscuit box standing in front of you.

(3) internally guided action inhibited through externally guided stop signal: e.g., stop teasing a classmate when a teacher suddenly appears.

(4) internally guided action inhibited through internally guided stop signal: e.g., resisting the impulse to cut in line.

This paper focuses only on when the inhibition is internally guided (i.e., only (2) and (4)). Formation of inhibition is analysed in four dimensions: intentionality (i.e., extent of intentionality involved), timing (i.e., point in time when the inhibition process is initiated), specificity (i.e., inhibition is selective or global), and the nature of the to be inhibited action (Ridderinkhof, van den Wildenberg, & Brass, 2014). According to this categorization for intentional inhibition, the extent of intentionality involved is still an open question. Furthermore, it is not exactly clear when intentional inhibition is initiated. Note that sufficient time is required to go through the intentional inhibition process. It is assumed that intentional inhibition is a relatively slow process (Ridderinkhof et al., 2014). Rather than intentionally inhibiting the action, the current action may be suspended first in order to buy time to decide whether to (re)initiate the action or not; cf. (Kühn & Brass, 2009). The third category specifically is relevant for cases when there are multiple action preparations in parallel. A question then is whether all the actions are going to be intentionally inhibited globally or only selective actions are intentionally inhibited. It seems that still there is no research on this category particular to intentional inhibition; in experiments always intentional inhibition for single action preparation has been addressed. The fourth category utilizes factors like temporal aspect (i.e., strong stimulus–response associations) and strength or prepotency of the action. Intention inhibition has been defined as a late veto mechanism (Filevich, Kühn, & Haggard, 2012; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012; Kühn, Haggard, & Brass, 2009; Brass & Haggard, 2008, 2007; Haggard, 2008). More importantly, in the intentional inhibition mechanism only an already initiated action will be inhibited, at the last possible moment.

There are more facts available from neuroscience evidence to differentiate the external inhibition from internal inhibition. Haggard (2008) has referred these two cases namely, early and late decisions about whether to act. Walsh et al. (2010) have noted in their research experiment on intentional inhibition that the intentional inhibition experience will occur some hundreds of milliseconds prior to voluntary actions which is in line with the past research evidences in (Libet, Gleason, Wright, & Pearl, 1983; Sirigu et al., 2004). There are criticisms on the hypotheses that
provide relevant brain areas which are contributing for intentional inhibition (Filevich, Kühn, & Haggard, 2012; Lynn, Muhle-Karbe, & Brass, 2014). Nevertheless, much research has confirmed that intentional inhibition is a core phenomenon in human brain. For example, according to (Kühn, Haggard, & Brass, 2009):

‘the current findings demonstrate the role of dorsal Fronto-median cortex (dFMC) in intentional inhibition of action, and its effective connectivity with areas involved in intention and preparation of action. We used a naturalistic task involving clear response affordances and impulsive actions, close to real-life experiences of action decision. dFMC was involved in intentional inhibition of such responses, in addition to intentional inhibition of self-generated actions reported previously. In accordance with a functional fractionation of intentional action by Brass and Haggard (Brass & Haggard, 2007), we could dissociate the area that is involved in the implementation of intentional inhibition (dFMC) from the area involved in the decision whether to act or not (RCZ). Our results represent a further step in addressing the question of self-control. Future research may benefit from taking a cognitive-motor approach to study clinical disorders of self-control.’ (Kühn, Haggard, & Brass, 2009, p. 2842)

Intentional inhibition is a phenomenon that is difficult to measure due to the lack of any behavioural output: how to distinguish an inhibited action from a never prepared action? The development of brain imaging and recording techniques and particular experimental set-ups (e.g., marble experiment or ramp task) have increased the attention of cognitive neuroscientist to explore this neglected branch in recent years (Kühn, Haggard, & Brass, 2009; Lynn, Muhle-Karbe, & Brass, 2014). The following are also in line with previous observations:

‘These results show that, in adults, intentional and externally guided inhibition activate a similar network of regions in lateral prefrontal regions (including grIFG) and ACC/preSMA. However, intentional inhibition also appears to be associated with distinct activation in dFMC, a region that is not involved in stimulus-driven inhibition (Brass & Haggard, 2007; Kühn et al., 2009; Schel et al., 2014). This dFMC region extending to the dorsal ACC is also found to be the main region important for intentional inhibition in a meta-analysis of fMRI studies looking at intentional inhibition (Filevich et al., 2012). Interestingly, although intentional inhibition cannot be easily manipulated because it is not triggered by an external signal or cue, it has been shown that varying the preceding context can influence both the likelihood of intentional inhibition and the level of activation of the dFMC during intentional inhibition (Schel et al., 2014). Thus, the underlying neural mechanisms of intentional inhibition appear to be at least partly distinct from the neural mechanisms underlying externally guided inhibition.’ (Schel, Scheres, & Crone, 2014, p. 240)

For intentional inhibition experiments most researchers focus on cancellation of motor responses through brain imaging and recording techniques for the related brain areas (e.g., dFMC). The dFMC area is considered to be the veto area that
generates endogenous top-down signals, which directly affect a current yet to execute action in order to stop or cancel it immediately. Nevertheless, there are criticisms on this, due to enabling the same areas on different inhibition processes (e.g., emotional inhibition that does not relate so clearly with any immediate action or behavioural output, but rather with face or body expression). Therefore, it is proposed that dFMC may play a role that detaches a person from its immediate urges (see Lynn, Muhle-Karbe, & Brass, 2014). Due to this, similar to other complex processes in the human brain intentional inhibition may also not directly relate to a particular area, but spread over many areas by cyclic loops. For example, it has been noted that areas like pre-SMA, IFC, and basal ganglia, which are mainly used in externally guided inhibition situations, also recruit in intentional inhibition (Ridderinkhof, van den Wildenberg, & Brass, 2014). As shown in literature such as (Filevich, Kühn, & Haggard, 2012; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012; Kühn, Haggard, & Brass, 2009; Brass & Haggard, 2008, 2007) decisions to suppress a prepared action may be based on a more complex process of intentional inhibition. It may well be the case that a primary valuation of the predicted effects of an action is considered positive so that executing the action seems justified, but in a wider context a negative valuation of predicted effects is obtained that still may lead to suppressing the action. Mostofsky and Simmonds (2008) have referred this twofold phenomenon (response selection and response inhibition) as the two sides of the same coin and provided experimental results. The Supplementary Motor Area (SMA) within the medial wall of the frontal cortex is shown to be responsible for response preparation, selection, and execution. The pre-Supplementary Motor Area (pre-SMA) is appeared to be responsible for response inhibition; (there is supportive evidence that early activation of neurons within the pre-SMA regions is involved in successful action selection) (Mostofsky & Simmonds, 2008). Mostofsky and Simmonds (2008) reviewed the findings for these from behavioral, brain imaging, and lesion effects studies; for example:

‘Based on the findings, we propose a theoretical construct in which response inhibition is viewed as a facet of response selection, such that response inhibition is an intentional process in which one actively selects to withhold a response while producing a goal oriented one (not moving). As with response selection, response inhibition therefore depends on medial frontal premotor circuits critical for motor response preparation, with variable roles of involvement of prefrontal circuits (as well as posterior cortical regions) necessary for guiding response inhibition based on the cognitive/social context of the task (i.e., ‘task demand’).’ (Mostofsky & Simmonds, 2008, p. 751)

In general, anterior regions of the brain are more involved in higher-order intentionality, whereas posterior regions contribute to generating physical action. A negative correlation has been observed between dFMC activation and primary
motor cortex activation when intentional inhibition occurs (Brass & Haggard, 2007). It has also been found that intentional inhibition can be unconsciously primed; this shows that conscious intentional inhibition is influenced by unconscious factors (Parkinson & Haggard, 2014). Furthermore, Ridderinkhof, Van den Wildenberg, and Brass (2014) have related ideomotor theory with motor control theories. In this process the classical stop-signal and go/nogo paradigm (e.g., see Van den Wildenberg et al., 2010) relates to sensorimotor inhibition, whereas anticipation of the sensory consequences of inhibition relates to ideomotor inhibition, which is more in line with intentional inhibition. Furthermore, they have observed that there is a large overlap in networks between sensorimotor inhibition and ideomotor inhibition, but with an additional recruitment of dFMC for ideomotor inhibition (Ridderinkhof, van den Wildenberg, & Brass, 2014). The following quotation provides more intuition on the interplay between sensorimotor and ideomotor inhibition:

‘Optimal action control can be viewed as a dynamic interplay between activation and suppression mechanisms that is achieved by an intricate and complex fronto-striatal circuitry (for review see Ridderinkhof, Forstmann, Wylie, Burle, & van den Wildenberg, 2011 and Wiecki & Frank, 2013). This action selection and override circuitry consist of tightly interconnected cortical regions and basal ganglia structures (most prominently the dorsolateral prefrontal cortex [PFC], the inferior frontal cortex [IFC], the pre-supplementary motor area [pre-SMA], the dorsal striatum, and the subthalamic nucleus [STN]). Sensorimotor and ideomotor action appears to be supported, in large measure, by shared neural mechanisms. However, there appear to be qualitative and quantitative differences as well. For instance, within the dorsal portions of the striatum, the posterior putamen and its associated premotor circuitry are thought to be more important for habitual sensorimotor action while the caudate and its associated PFC circuitry are more central to goal-directed ideomotor action (de Wit et al., 2012; Tanaka, Balleine, & O’Doherty, 2008; Tricomi, Balleine, & O’Doherty, 2009; Balleine & O’Doherty, 2010). The engagement of the pre-SMA appears to be greater when prepotent sensorimotor actions conflict with goal-directed ideomotor actions. Similarly, we may ask whether sensorimotor and ideomotor inhibition also rests on overlapping or separable neural mechanisms. A vast literature documents how brain systems inhibit or override responses that are not instrumental to our current goals (for extensive review see Aron & Poldrack, 2006, Aron, Robbins, & Poldrack, 2014 and Ridderinkhof et al., 2011). This literature capitalizes on situations where inhibition was called forth by extraneous cues (such as stop signals) or by action conflicts (e.g., Forstmann et al., 2008). The patterns that emerge from these reviews suggest that the dorsolateral PFC provides top-down guidance to action selection areas, the pre-SMA engages response inhibition as an instrument of action selection, the IFC (mostly in the right hemisphere) is recruited to aid in implementing response inhibition in more demanding situations, and the basal ganglia keep all responses in check until the final call is received from upstream. Direct connections from the pre-SMA and IFG to basal ganglia structures (most prominently the anterior dorsal striatum and the STN) serve to keep basal ganglia output in check until
ideomotor action selection has completed (Forstmann et al., 2008; Herz et al., 2014; Jahfari et al., 2011, 2012). (Ridderinkhof, van den Wildenberg, & Brass, 2014, p. 260)

The importance of intentional inhibition for healthy social functioning in a society has been highlighted in (Mostofsky & Simmonds, 2008; Cohen, Berkman, & Lieberman, 2013). They provide literature about clinical disorders with impairments in such control, like Anarchic Hand Syndrome (AHS) (Della Sala & Marchetti, 2005; Kritikos, Breen, & Mattingley, 2005), Attention Deficit Hyperactivity Disorder (ADHD) (Barkley, 1997; Roberts, Fillmore, & Milich, 2011), substance abuse, and pathological gambling. According to Cohen, Berkman, and Lieberman (2013) self-control is a wider spectrum and intentional inhibition is a key phenomenon:

'A good general definition is that self-control is “the overriding or inhibiting of automatic, habitual, or innate behaviors, urges, emotions, or desires that would otherwise interfere with goal directed behavior”’ (Muraven, Shmueli, & Burkley, 2006, p. 524). As this definition indicates, many different methods can be used to study self-control, ranging from inhibiting a motor response to regulating an emotion to suppressing the temptation to eat sweets. In addition to these explicit, intentional forms of self-control, it is possible to exert control without an explicit goal to do so (i.e., automatically or incidentally) given the right situation. For example, in priming paradigms, participants are not explicitly aware that they saw a prime, but the implicit encoding of primes can cause incidental behavioral control. Additionally, it is possible to implicitly or incidentally regulate an affective response without awareness (for a review, see Berkman & Lieberman, 2009).’ (Cohen, Berkman, & Lieberman, 2013, p. 417)

Humans are always responsible for their actions as they could have decided to withhold them with their innate ability of intentional inhibition, which has been contributed to differentiate us from other species. In the current paper this more complex process involving intention inhibition is modelled.

3.3. Description of the Computational Model

An overview of the postulated computational cognitive agent model is presented in Fig. 1 below; the state labels used in the model are summarized in Table 1. The proposed model consists mainly two loops namely impact prediction loop (Filevich, Kühn, & Haggard, 2012), and effect prediction loop (as-if body loop) (Damasio, 1999, 2005). The impact prediction loop models effects of awareness states while the effect prediction loop mainly demonstrates unconscious behavior. In this model Prior Ownership and Retrospective Ownership states are considered unconscious ownership states and the Prior Awareness and Retrospective Awareness states as conscious ownership states. Therefore, this model contributes for both of conscious and unconscious aspects of the processes. More details of the model construction can be found in the subsections following below.
3.3.1. Inputs to the model and their representation

The model uses three inputs: stimulus \( s \), context \( c \), and effect \( b \). The stimulus \( s \) is a state associated with a detectable change in the bodily (e.g., self-generated facial expression) or external (e.g., emotional state of another person) environment that may lead to execute an action. A context \( c \) is of a more general type of perceived information, not directly relating to an action execution. As an example, observing a fire on two collided vehicles can be considered a context under the assumption that agent will not react on it. However, noticing a person who has been trapped in a vehicle which is on fire is considered a stimulus under the assumption...
that the agent will do something on this. Context \( c \) can be of different types: when the context \( c \) is self, that indicates a focus on self, whereas for context \( c \) an observed agent B (other), indicates on another person B. Effect \( b \) is the input to the model indicating effect of execution of an action \( a \). The model has the ability to develop relevant prior and retrospective states and the communication of ownership.

Inputs that are coming from world states: WS(\( s \)), WS(\( c \)), and WS(\( b \)); lead to Sensor States: SS(\( s \)), SS(\( c \)), and SS(\( b \)) respectively. Sensor states lead to further internal processes according to the following causal sequence (Treur, 2013) (which is referred as body loop in Fig. 1):

\[
\text{sensing a stimulus} \rightarrow \text{sensory representation of stimulus} \rightarrow \text{preparation for bodily response} \rightarrow \text{sensing the bodily response} \rightarrow \text{sensory representation of the bodily response} \rightarrow \text{feeling the emotion}
\]

In practical context these inputs can be assigned to a necessary situation that may lead to many options (i.e., PA(\( a \))) and through the initial value given to the WS(\( s \)) and WS(\( c \)) states (value 1 means the strongest input and value 0 means no input), the model initiates its behaviour. For an example, assume that someone is asking a question in a conference, where that question can be considered as a stimulus \( s \) and the conference situation can be considered as the context \( c \).

**3.3.2. Effect prediction loop**

Effect prediction or as-if body loop is a form of the casual sequence stated in Section 3.1. It moves from preparation for bodily response to sensory representation of the bodily response to feeling the associated emotion (Damasio, 1999, 2005). Through the effect prediction loop it demonstrates mental formation of action selection (or rejection) in a parallel fashion: PA(\( a_i \)) where \( i \) goes from 1 to \( n \).

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WS(W)</td>
<td>world state W (W is a context ( c ), stimulus ( s ), or effect ( b ))</td>
</tr>
<tr>
<td>SS(W)</td>
<td>sensor state for W</td>
</tr>
<tr>
<td>SR(W)</td>
<td>sensory representation of W</td>
</tr>
<tr>
<td>PA(( a ))</td>
<td>preparation for action ( a ): action ( a ) can be either objective (( obj )) or subjective (( sub ))</td>
</tr>
<tr>
<td>F(( b ))</td>
<td>feelings of action ( a ) after: effect prediction (( obj )) or impact prediction (( sub )) or action execution (( obj ))</td>
</tr>
<tr>
<td>PO(( a, b, c, s ))</td>
<td>prior ownership state for action ( a ) with ( b, c, ) and ( s )</td>
</tr>
<tr>
<td>RO(( a, b, c, s ))</td>
<td>retrospective ownership state for ( a ) with ( b, c, ) and ( s )</td>
</tr>
<tr>
<td>EA(( a ))</td>
<td>execution of action ( a )</td>
</tr>
<tr>
<td>EO(( a, b, c, s ))</td>
<td>communication of ownership of ( a ) with ( b, c, ) and ( s )</td>
</tr>
<tr>
<td>PAwr(( a, b, c, s ))</td>
<td>prior-awareness state for action ( a ) with ( b, c, ) and ( s )</td>
</tr>
<tr>
<td>RAwr(( a, b, c, s ))</td>
<td>retrospective-awareness state for action ( a ) with ( b, c, ) and ( s )</td>
</tr>
<tr>
<td>PD(( b ))</td>
<td>performative desires for ( b )</td>
</tr>
<tr>
<td>CD(( b ))</td>
<td>constitutive desires for ( b ): the desires to be in control of one’s self</td>
</tr>
</tbody>
</table>

**Table 1. Nomenclature for Fig. 1.**
Therefore it is possible for an agent to evaluate in parallel many possible actions \( PA(a_i) \) for a triggered stimulus \( s \) without actually executing any of those. Finally the best valued candidate will be selected for execution and other choices naturally get diluted. Furthermore, this model takes the influence of performative desires for preparation of action. Performative desires are mainly contributing for short term interests/goals that influence either selecting or rejecting due to satisfactory or less satisfactory valuation of an action (Filevich, Kühn, & Haggard, 2012). This rejection is different from intentional inhibition in the literature (Filevich, Kühn, & Haggard, 2012; Zhang, Hughes, & Rowe, 2012; van den Wildenberg et al., 2010). These desires are rapidly changing and having a relatively low lifespan when comparing with constitutive desires (i.e., which drives a person through his or her long term driven aspirations) explained in the next section. For an example to explain the performative desires: when people enjoy (regularly) burgers and soda, they mainly satisfy their performative desires. As another example; assume that in a conference a gentlemen asked you a question which you feel as an offending question triggering anger (based on performative desires); nevertheless as a professional you will answer that question very politely and calmly in front of all the audience (due to the effect of your constitutive desires). In this effect prediction process the main driving force is formed by the objective preparations for action options: \( PA(a_{obj}) \). As these options are driven by external inputs and short term interests/goals (performative desires) they have more objective features than subjective features. These objective features relate to habits or stereotypes. Here it is assumed that more unconscious aspects relate to objective features whereas more conscious aspects relate to subjective features. Therefore, effect prediction loops provide unconscious influences through objective action preparations whereas in impact prediction loops provide conscious influences through subjective features together with constitutive desires (see Filevich, Kühn, & Haggard, 2012).

In the effect prediction loop, the preparation for action in objective terms \( PA(a_{obj}) \) gets effects from the sensory representation of context \( SR(c) \), the sensory representation of stimulus \( SR(s) \), the performative desires \( PD(b) \) of \( b \), and the feeling of effect prediction \( F(b_{obj}) \) of action \( a_{obj} \). The sensory representation \( SR(b_{obj}) \) of effect \( b_{obj} \) gets effects from the preparation for action in objective terms \( PA(a_{obj}) \), the sensory state \( SS(b) \) of executed action of \( a \), and the prior ownership state for action \( a \) with \( b \), \( c \), and \( s \). Here, the effect from prior ownership to sensory representation of effect \( b_{obj} \) is a suppression, which contributes to dilute the state level after the action gets executed. The feeling of effect prediction \( F(b_{obj}) \) of action \( a_{obj} \) gets effects from the sensory representation \( SR(b_{obj}) \) of effect \( b_{obj} \). Performative desires \( PD(b) \) of \( b \) gets effects from the sensory representation of context \( SR(c) \), the sensory representation of stimulus \( SR(s) \), and the sensory representation \( SR(b_{obj}) \) of effect \( b_{obj} \).
3.3.3. Body loop or action execution

The body loop is the causal sequence presented in Section 3.1. Once the model isolated the best valued candidate through the as-if body loop, the selected action will be get executed through the body loop. The effect from sensory representation of effect $h_{obj}$ on performative desires is suppression, which contributes to dilute the performative desires from the effects of action execution for actions satisfying the desire.

3.3.4. Impact prediction loop

The impact prediction loop is a loop parallel to the as-if body loop and drives the person through his or her long term driven aspirations (referred here as constitutive desires). Constitutive desires are more towards the subjective aspects. This makes an additional contribution to the current action selection process by considering the longer term consequences and implications of the considered action (Balleine & O’Doherty, 2010; de Wit et al., 2012; Tanaka, Balleine, & O’Doherty, 2008; Tricomi, Balleine, & O’Doherty, 2009). Through this the person can demonstrate its maturity as a ‘social entity’ instead of merely putting his or her laymen feelings into actions (Mostofsky & Simmonds, 2008; Cohen, Berkman, & Lieberman, 2013). Using the impact prediction loop based on the constitutive desires, the person has the ability to decide whether it is required to abandon the action selection based on performative desires: intentional inhibition (Filevich, Kühn, & Haggard, 2012; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012; Kühn, Haggard, & Brass, 2009; Brass & Haggard, 2008, 2007). Intentional inhibition does not simply reset or drive the performative desires to a non-action situation. This phenomenon is fundamentally different from an action selection process in which non-action is simply another alternative (Filevich, Kühn, & Haggard, 2012) (unintentional inhibition occurs prior to conscious awareness).

In this loop the preparation for abandoning the action in subjective terms PA($a_{sub}$) gets effects from the sensory representation of stimulus SR($s$), the feeling of effect prediction $F(h_{obj})$ of action $a_{obj}$, the constitutive desires CD($b$) of $b$, the feeling of effect prediction $F(h_{sub})$ for action $a_{sub}$, the prior awareness state PAwr($a$, $b$, $c$, $s$) for action $a$ with $b$, $c$, and $s$, and the retrospective awareness state RAwr($a$, $b$, $c$, $s$) for action $a$ with $b$, $c$, and $s$. Having effects from prior and retrospective awareness to the preparation for abandoning the action in subjective terms, models the idea of impact prediction loop taking into account more conscious elements in the process (Pacherie, 2008). The sensory representation SR($h_{sub}$) of effect $h_{sub}$ gets effects from the preparation for action in subjective terms PA($a_{sub}$). The feeling of effect prediction $F(h_{sub})$ of action $a_{sub}$ gets effects from the sensory representation SR($h_{sub}$) of effect $h_{sub}$, and constitutive desires CD($b$). When this feeling has a high
activation level, this means that the action does not satisfy the constitutive desire CD(b), so preparation PA(a_{sub}) to abandon the action should become high as well. Constitutive desires CD(b) of b themselves get effects from the sensory representation of stimulus (SR(s)), and the sensory representation SR(b_{obj}) of effect b_{obj}. Here, the effect from sensory representation of effect b_{obj} to constitutive desires is suppression, which contributes to dilute the constitutive desires from the effects of action execution for actions satisfying this desire. Note that actions that do satisfy the constitutive desires may well be different from the actions satisfying the performative desires. For example, taking fast food may satisfy the performative desires whereas taking healthy food may satisfy the constitutive desires. Indeed, such differences may serve as triggers for intentional inhibition.

3.3.5. Ownership states

An important function of ownership states is that they mainly determine to what extent an agent attributes an action to himself or to another agent. In addition, they play their role in decisions to execute actions. This model adopts parts of the ownership-related states from cognitive agent model presented in (Treur, 2012). There are two ownership states: the prior ownership state PO(a, b, c, s) for action a with b, c, and s, and the retrospective ownership state RO(a, b, c, s) for a with b, c, and s. These ownership states are mainly assumed to be unconscious ownership states. Prior ownership state emerges prior to the action execution whereas retrospective ownership state emerges precede to the action execution.

The prior ownership state PO(a, b, c, s) for action a with b, c, and s gets effects from the sensory representation of context SR(c), the feeling of effect prediction F(b_{obj}) of action a_{obj}, the preparation for action in objective terms PA(a_{obj}), and the retrospective ownership state RO(a, b, c, s) for action a with b, c, and s. Here, the effect from the retrospective ownership state for action a with b, c, and s is a suppression, which contributes to dilute the prior awareness once the retrospectives awareness was developed. The retrospective ownership state RO(a, b, c, s) for action a with b, c, and s gets effects from the execution EA(a) of action a, the feeling F(b_{obj}) of the effect prediction of action a_{obj}, the prior ownership state PO(a, b, c, s) for action a with b, c, and s, the sensory representation of context SR(c), and the preparation for action in subjective terms PA(a_{sub}). Here, the effect from preparation for action in subjective terms to retrospective ownership is a suppression link.

3.3.6. Awareness states

This model adopts parts of the awareness related states from the cognitive agent model presented in (Thilakarathne & Treur, 2013b, 2015). Awareness mainly concerns a person’s subjective experience of a cognitive content. There are two
awareness states: the prior awareness state \( \text{PAwr}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \), and the retrospective awareness state \( \text{RAwr}(a, b, c, s) \) for \( a \) with \( b, c \), and \( s \). These awareness states are mainly assumed to be conscious ownership states. Prior awareness state precedes to the prior ownership state and retrospective awareness precedes to the retrospective ownership state.

The prior awareness state \( \text{RAwr}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \) gets effects from the feeling of effect prediction \( F(b_{sub}) \) of action \( a_{sub} \), the preparation for action in subjective terms \( \text{PA}(a_{sub}) \), the prior ownership state \( \text{PO}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \), the feeling \( F(b_{obj}) \) of effect prediction of action \( a_{obj} \), and the retrospective awareness state \( \text{RAwr}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \). Lack of awareness of experienced feelings has shown that it may make it difficult to select the appropriate action selection according to the situation (Barrett, Gross, Christensen, & Benvenuto, 2001). Here, the effect from retrospective awareness to prior awareness is a suppression link; so that once the retrospective awareness developed it leads to suppress the prior awareness. The retrospective awareness state \( \text{RAwr}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \) gets effects from the feeling \( F(b_{sub}) \) of effect prediction of action \( a_{sub} \), the prior awareness state \( \text{PAwr}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \), the retrospective ownership state \( \text{RO}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \), and the feeling \( F(b_{obj}) \) of effect prediction of action \( a_{obj} \).

### 3.3.7. Execution and communication states

The execution state \( \text{EA}(a) \) of action \( a \) is the state which finally performs the action which was internally formed and decided upon. The main feature of this model is the causal relation from preparation for action in subjective terms \( \text{PA}(a_{sub}) \) to execution of action: intention inhibition (Filevich, Kühn, & Haggard, 2012; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012; Kühn, Haggard, & Brass, 2009). Execution \( \text{EA}(a) \) of action \( a \) gets effects from the prior awareness state \( \text{PAwr}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \), the preparation for action in subjective terms \( \text{PA}(a_{sub}) \), the prior ownership state \( \text{PO}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \), and the preparation for action in objective terms \( \text{PA}(a_{obj}) \).

Communication \( \text{EO}(a, b, c, s) \) of ownership of \( a \) with \( b, c \), and \( s \) is based on the capability of expressing the ownership of a selected action (where \( c \) can be either other or self). Communication \( \text{EO}(a, b, c, s) \) of ownership of \( a \) with \( b, c \), and \( s \) is affected by the retrospective awareness state \( \text{RAwr}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \), and the retrospective ownership state \( \text{RO}(a, b, c, s) \) for action \( a \) with \( b, c \), and \( s \). Through this state, this model can be used in a social context with multiple agents and each can express their details of ownership with the awareness. Therefore, the current model has options for further expansions to include even joint decision making processes.
3.3.8. Dynamics of model compilation

Connections between state properties (the arrows in Fig. 1) have weights \( w_k \), as indicated in Table 2. In this table the column LP refers to the (temporally) Local Properties LP1 to LP17. Formal specification in hybrid LEADSTO format (Bosse, Jonker, Van Der Meij, & Treur, 2007) is shown in Table 3 for LPs. LEADSTO is a hybrid modeling language in which a dynamic property or temporal causal relation \( a \rightarrow b \) denotes that when a state property \( a \) (or conjunction thereof) occurs, then after a certain time delay, state property \( b \) will occur. The time delay defined in LEADSTO is taken as a uniform time step \( \Delta t \) here. A weight \( w_k \) has a value between -1 and 1 and may depend on the specific context \( c \), stimulus \( s \), action \( a \) and/or effect state \( b \) involved. By varying these connection strengths, different possibilities for the characteristics and repertoire offered by the modeled person can be realised. Note that usually weights are assumed non-negative, except for the inhibiting connections, such as \( w_1, w_3, w_{11}, w_{25}, w_{27}, w_{35}, w_{40}, w_{42}, \) and \( w_{45} \).

The dynamics following the connections between the states in Fig. 1 has been

<table>
<thead>
<tr>
<th>from state</th>
<th>to state</th>
<th>weights</th>
<th>LP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( EA(a) ), ( EA(a) ), ( EO(a,b,c,s) )</td>
<td>( WS(W) )</td>
<td>( w_1, w_2, w_3 )</td>
<td>LP1</td>
</tr>
<tr>
<td>( WS(s) ), ( WS(c) ), ( WS(b) )</td>
<td>( SS(W) )</td>
<td>( w_4, w_5, w_6 )</td>
<td>LP2</td>
</tr>
<tr>
<td>( SS(s) ), ( SS(c) )</td>
<td>( SR(W) )</td>
<td>( w_7, w_8 )</td>
<td>LP3</td>
</tr>
<tr>
<td>( SS(b) ), ( PA(a_{obj}) ), ( PO(a,b,c,s) )</td>
<td>( SR(b_{obj}) )</td>
<td>( w_9, w_{10}, w_{11} )</td>
<td>LP4</td>
</tr>
<tr>
<td>( PA(a_{sub}) )</td>
<td>( SR(b_{sub}) )</td>
<td>( w_{12} )</td>
<td>LP5</td>
</tr>
<tr>
<td>( SR(c) ), ( SR(s) ), ( PD(b) ), ( F(b_{obj}) )</td>
<td>( PA(a_{obj}) )</td>
<td>( w_{13}, w_{14}, w_{15}, w_{16} )</td>
<td>LP6</td>
</tr>
<tr>
<td>( F(b_{sub}) ), ( CD(b) ), ( SR(s) ), ( F(b_{obj}) ), ( RAwr(a,b,c,s) ), ( PAwr(a,b,c,s) )</td>
<td>( PA(a_{sub}) )</td>
<td>( w_{17}, w_{18}, w_{19}, w_{20}, w_{21}, w_{22} )</td>
<td>LP7</td>
</tr>
<tr>
<td>( SR(c) ), ( SR(s) ), ( SR(b_{obj}) )</td>
<td>( PD(b) )</td>
<td>( w_{23}, w_{24}, w_{25} )</td>
<td>LP8</td>
</tr>
<tr>
<td>( SR(s) ), ( SR(b_{obj}) )</td>
<td>( CD(b) )</td>
<td>( w_{26}, w_{27} )</td>
<td>LP9</td>
</tr>
<tr>
<td>( SR(b_{obj}) ), ( PD(b) )</td>
<td>( F(b_{obj}) )</td>
<td>( w_{28}, w_{29} )</td>
<td>LP10</td>
</tr>
<tr>
<td>( SR(b_{sub}) ), ( CD(b) )</td>
<td>( F(b_{sub}) )</td>
<td>( w_{30}, w_{31} )</td>
<td>LP11</td>
</tr>
<tr>
<td>( SR(c) ), ( F(b_{obj}) ), ( PA(a_{obj}) ), ( RO(a,b,c,s) )</td>
<td>( PO(a,b,c,s) )</td>
<td>( w_{32}, w_{33}, w_{34}, w_{35} )</td>
<td>LP12</td>
</tr>
<tr>
<td>( F(b_{sub}) ), ( PA(a_{sub}) ), ( PO(a,b,c,s) ), ( F(b_{obj}) ), ( RAwr(a,b,c,s) ), ( PAwr(a,b,c,s) )</td>
<td>( PAwr(a,b,c,s) )</td>
<td>( w_{36}, w_{37}, w_{38}, w_{39}, w_{40} )</td>
<td>LP13</td>
</tr>
<tr>
<td>( PAwr(a,b,c,s) ), ( PA(a_{obj}) ), ( PO(a,b,c,s) ), ( PA(a_{sub}) )</td>
<td>( EA(a) )</td>
<td>( w_{41}, w_{42}, w_{43}, w_{44} )</td>
<td>LP14</td>
</tr>
<tr>
<td>( PA(a_{sub}) ), ( SR(c) ), ( PO(a,b,c,s) ), ( F(b_{obj}) ), ( EA(a) )</td>
<td>( RO(a,b,c,s) )</td>
<td>( w_{45}, w_{46}, w_{47}, w_{48}, w_{49} )</td>
<td>LP15</td>
</tr>
<tr>
<td>( F(b_{sub}) ), ( PAwr(a,b,c,s) ), ( RO(a,b,c,s) ), ( F(b_{obj}) )</td>
<td>( RAwr(a,b,c,s) )</td>
<td>( w_{50}, w_{51}, w_{52}, w_{53} )</td>
<td>LP16</td>
</tr>
<tr>
<td>( RAwr(a,b,c,s) ), ( RO(a,b,c,s) )</td>
<td>( EO(a,b,c,s) )</td>
<td>( w_{54}, w_{55} )</td>
<td>LP17</td>
</tr>
</tbody>
</table>
designed based on a dynamical systems perspective; e.g., (Bosse, Jonker, Van Der Meij, & Treur, 2007). During processing, each state property has a strength represented by a real number between 0 and 1; variables V (possibly with subscripts) run over these values. In dynamic property specifications, this is added as a last argument to the state property expressions. Below, \( f \) is a function for which different choices can be made. In the example simulations, for the states that are affected by only one state (i.e., in LP1, LP2, and LP3), the function \( f \) is taken as the identity function \( f(W) = W \), and for the other states \( f \) is a combination function based on the logistic threshold function as in equations (1), (2) and (3). In these equations, \( \sigma \) is the steepness and \( \tau \) the threshold; these are configuration parameters that change the shape of the curve and its midpoint on the X-axis.

\[
f(X) = \text{th}(\sigma, \tau, X) \quad \text{ when } X \geq 0
\]

where \( X = \sum_j \omega_{ji} y_j \) and \( \omega_{ji} \) represents the incoming links

for state \( i \) and \( y_j \) represents the strength of state \( j \)

\[
\text{th}(\sigma, \tau, X) = \left( \frac{1}{1 + e^{-\sigma(X-\tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{-\sigma\tau})
\]

\( f(X) = 0 \quad \text{ when } X < 0 \) (3)

Activation of a state depends on multiple other states that are directly attached to it; therefore incoming activation levels from other states are combined to some aggregated input and affect the activation level according to a specification in differential equation format, as in equation (4) (where \( y_i \) is the activation level of state \( i \)).

\[
\frac{dy_i}{dt} = \gamma_i \left[ \text{th} \left( \sigma, \tau, \sum_j \omega_{ji} y_j \right) - y_i \right]
\]

Parameter \( \gamma \) is an update speed factor, indicating the speed by which an activation level is updated upon received input from other states. In this model two speed factor values are used: one for the internal states (states which are inside the dotted box in Fig. 1), and the other one for the external states. The internal states’ speed factor is higher than the external states (adhering to the phenomenon that brain neurons are activating much faster than sensor and effector organs). To obtain a computational specification for temporal simulation of each state, a difference equation is used in the form of equation (5).

\[
y_i(t + \Delta t) = y_i(t) + \gamma_i \left[ \text{th} \left( \sigma, \tau, \sum_{j \in s(t)} \omega_{ji} y_j \right) - y_i(t) \right] \Delta t
\]
By having different values for each parameter (i.e., for weight values $\omega_i$, time step size $\Delta t$, slow and fast speed factors $\gamma$, steepness $\sigma_i$, threshold $\tau_i$) the agent can facilitate a wide variety of behaviours. Each LP in Table 3 is represented in a computational form and the dynamics of the system is achieved through evaluating the causal effects through a set of difference equations as in equation (5). For each discrete time step $\Delta t$ the behaviour of each state is calculated; in this way the emergence of the behaviour is traced based on an identified parameter value set. From a mathematical point of view, the dynamics of the model is (numerically)

Table 3. Formal specification of the model in LEADSTO format.

<table>
<thead>
<tr>
<th>LP</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP1</td>
<td>$\text{EA}(a, V_1) &amp; \text{WS}(b, V_2) \implies \text{WS}(b, V_2 + \gamma [f(w_2 V_1) - V_2]) \Delta t$</td>
</tr>
<tr>
<td>LP2</td>
<td>$\text{WS}(s, V_1) &amp; \text{SS}(s, V_2) \implies \text{SS}(s, V_2 + \gamma [f(w_4 V_1) - V_2]) \Delta t$</td>
</tr>
<tr>
<td>LP3</td>
<td>$\text{SS}(s, V_1) &amp; \text{SR}(s, V_2) \implies \text{SR}(s, V_2 + \gamma [f(w_5 V_1) - V_2]) \Delta t$</td>
</tr>
<tr>
<td>LP4</td>
<td>$\text{SR}(b_{obj}, V_1) &amp; \text{PA}(a_{obj}, V_2) &amp; \text{PO}(a, b, c, s, V_3) &amp; \text{SR}(b_{obj}, V_4) \implies \text{SR}(b_{obj}, V_4 + \gamma [f(w_9 V_1, w_{10} V_2, w_{11} V_1 - V_4]) \Delta t)$</td>
</tr>
<tr>
<td>LP5</td>
<td>$\text{PA}(a_{sub}, V_1) &amp; \text{SR}(b_{sub}, V_2) \implies \text{SR}(b_{sub}, V_2 + \gamma [f(w_{12} V_1) - V_2]) \Delta t)$</td>
</tr>
<tr>
<td>LP6</td>
<td>$\text{SR}(c, V_1) &amp; \text{SR}(s, V_2) &amp; \text{PD}(b, V_3) &amp; \text{F}(b_{obj}, V_4) &amp; \text{PA}(a_{obj}, V_3) \implies \text{PA}(a_{obj}, V_5 + \gamma [f(w_{13} V_1, w_{14} V_2, w_{15} V_2, w_{16} V_4 - V_5]) \Delta t)$</td>
</tr>
<tr>
<td>LP7</td>
<td>$\text{F}(b_{sub}, V_1) &amp; \text{CD}(b, V_2) &amp; \text{SR}(s, V_3) &amp; \text{F}(b_{obj}, V_4) &amp; \text{PA}(a_{obj}, V_3) \implies \text{PA}(a_{obj}, V_3 + \gamma [f(w_{17} V_1, w_{18} V_2, w_{19} V_3, w_{20} V_4, w_{21} V_5, w_{22} V_6 - V_7]) \Delta t)$</td>
</tr>
<tr>
<td>LP8</td>
<td>$\text{SR}(s, V_1) &amp; \text{SR}(s, V_2) &amp; \text{SR}(b_{obj}, V_3) &amp; \text{PD}(b, V_4) \implies \text{PD}(b, V_4 + \gamma [f(w_{23} V_1, w_{24} V_2, w_{25} V_3 - V_4]) \Delta t)$</td>
</tr>
<tr>
<td>LP9</td>
<td>$\text{SR}(b_{obj}, V_1) &amp; \text{PD}(b, V_2) &amp; \text{F}(b_{obj}, V_3) \implies \text{F}(b_{obj}, V_3 + \gamma [f(w_{26} V_1, w_{27} V_2 - V_3]) \Delta t)$</td>
</tr>
<tr>
<td>LP10</td>
<td>$\text{SR}(b_{sub}, V_1) &amp; \text{PD}(b, V_2) &amp; \text{F}(b_{sub}, V_3) \implies \text{F}(b_{sub}, V_3 + \gamma [f(w_{30} V_1, w_{31} V_2 - V_3]) \Delta t)$</td>
</tr>
<tr>
<td>LP11</td>
<td>$\text{SR}(b_{sub}, V_1) &amp; \text{CD}(b, V_2) &amp; \text{F}(b_{sub}, V_3) \implies \text{F}(b_{sub}, V_3 + \gamma [f(w_{32} V_1, w_{33} V_2, w_{34} V_3, w_{35} V_4 - V_3]) \Delta t)$</td>
</tr>
<tr>
<td>LP12</td>
<td>$\text{SR}(c, V_1) &amp; \text{F}(b_{obj}, V_2) &amp; \text{PA}(a_{obj}, V_3) &amp; \text{RO}(a, b, c, s, V_4) &amp; \text{PO}(a, b, c, s, V_5) \implies \text{PO}(a, b, c, s, V_5 + \gamma [f(w_{36} V_1, w_{37} V_2, w_{38} V_3, w_{39} V_4, w_{40} V_3 - V_6]) \Delta t)$</td>
</tr>
<tr>
<td>LP13</td>
<td>$\text{PA}(a_{sub}, V_1) &amp; \text{PO}(a_{obj}, V_3) &amp; \text{RO}(a, b, c, s, V_4) &amp; \text{PA}(a_{obj}, V_3) \implies \text{PA}(a_{obj}, V_3 + \gamma [f(w_{41} V_1, w_{42} V_2, w_{43} V_3, w_{44} V_4 - V_5]) \Delta t)$</td>
</tr>
<tr>
<td>LP14</td>
<td>$\text{PA}(a_{sub}, V_1) &amp; \text{SR}(c, V_2) &amp; \text{RO}(a, b, c, s, V_5) \implies \text{RO}(a, b, c, s, V_6 + \gamma [f(w_{45} V_1, w_{46} V_2, w_{47} V_3, w_{48} V_4, w_{49} V_5 - V_6]) \Delta t)$</td>
</tr>
<tr>
<td>LP15</td>
<td>$\text{PA}(a_{sub}, V_1) &amp; \text{SR}(c, V_2) \implies \text{RO}(a, b, c, s, V_3) \implies \text{RO}(a, b, c, s, V_3 + \gamma [f(w_{50} V_1, w_{51} V_2, w_{52} V_3, w_{53} V_4 - V_5]) \Delta t)$</td>
</tr>
<tr>
<td>LP16</td>
<td>$\text{F}(b_{sub}, V_1) &amp; \text{PA}(a_{obj}, V_3) &amp; \text{PO}(a, b, c, s, V_4) &amp; \text{F}(b_{obj}, V_4) \implies \text{F}(b_{obj}, V_4 + \gamma [f(w_{54} V_1, w_{55} V_2, w_{56} V_3, w_{57} V_4 - V_5]) \Delta t)$</td>
</tr>
<tr>
<td>LP17</td>
<td>$\text{PA}(a_{obj}, V_1) &amp; \text{RO}(a, b, c, s, V_2) &amp; \text{EO}(a, b, c, s, V_3) \implies \text{EO}(a, b, c, s, V_3 + \gamma [f(w_{58} V_1, w_{59} V_2 - V_3]) \Delta t)$</td>
</tr>
</tbody>
</table>
solving the differential equations of LP1 to LP17 by assuming that at time \( t=0 \), \( WS(s) \) and \( WS(c) \) holds value 1 as activation level.

### 3.4. Analysis of the Model Based on Simulation Experiments

This section discusses five simulation experiments to analyse the functionality of the designed model in different scenarios. In the first scenario a stimulus and context lead for a prepared action that has satisfactory predicted effects, but nevertheless the action is not performed due to intentional inhibition. The second scenario shows the backward compatibility of the model with (Thilakarathne & Treur, 2013b) which does not have the impact prediction. Scenario three is identical to the first but intentional inhibition does not occur and the action is performed. The fourth scenario shows a case behavior in a totally unconscious mode due to poor connections for the action effect prediction loop and the awareness states. Finally in the fifth scenario a situation is considered where the person has some cognitive impairment (or difficulty) in activating constitutive desires, together with subjective action preparation: \( PA(a_{sub}) \).

#### 3.4.1. Selecting values for the connection weights

Selecting suitable weight values for connections for neurological and behavioral agent models is a specific nontrivial issue. To identify suitable weight values for the connections in the current model an analytically driven approach was used (see Thilakarathne & Treur, 2015). First, a set of scenarios was identified for which the outcome can be at least partially identified in advance based on neurological and behavioral evidence from the literature. Based on that heuristic knowledge for a selected scenario the weight values were calibrated to simulate a pattern as expected. Once a set of values for these parameters were obtained for the selected scenario, the same values were used for another scenario, for which also an expected pattern was identified.

This new scenario requires changes to some parameters in order to adapt to it: for example, scenario 2 differs from scenario 1 only by setting values the values of \( w_{19}, w_{20}, w_{21}, w_{22}, w_{26}, \) and \( w_{27} \) to 0. If the previously identified values provide simulation results for the new scenario as expected, then the previously obtained parameter values become more justified. If, in contrast, the simulation results for the new scenario are not as expected, it is required to change some of the selected parameter values (based on the sensitivity of certain parameters on the required final output) until the simulations for this new scenario give results as expected.

If for a new scenario there are any changes to the previously obtained parameter values, then all previously addressed scenarios should be re-addressed. Through this iterative process after a number of cycles through a number of scenarios it is possible to obtain set of parameters values that is suitable for all these scenarios.
During this iteration process, if it is performed smoothly, the required changes for the parameter values will get lower and lower over time, according to a converging process.

Through the above mentioned approach the parameter values in Table 4 were obtained for the connections in the model. Threshold (τ) and steepness (σ) values used for scenarios have been listed in Table 5. Furthermore; the step size (Δt) taken is 0.25. The slow value 0.5 for γ was applied for external processes modeled by LP1, and LP2, and the fast value 0.9 for γ for the internal processes modeled by the other LP’s. The simulation results given in subsequent subsections are depicted using two axes: time and activation level. In this representation, time is represented as a series of step size (Δt) units rather than mapping with actual processing time. Being Δt is a parameter it is a decision choice in this model whether how small or large this value should be; nevertheless with more empirical evidences it may be better to represent actual processing time. By having a logistic threshold function as in equation (2) the value of the activation levels (the Y axis of the simulation graphs) is bounded between 0 and 1.

Table 4. Connection weight values used for cognitive agent model (note: all blank cells hold the respective value of S1 cell and further where S stands for Scenario).

<table>
<thead>
<tr>
<th>Weights</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>Weights</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
</tr>
</thead>
<tbody>
<tr>
<td>w1</td>
<td>-0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w26-29</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w2</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w30</td>
<td>0.9</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w3</td>
<td>-0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w31</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w32-34</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w5-6</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w35</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w7</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w36-37</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w8-9</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w38</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>w10</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w39</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>w11</td>
<td>-0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w40</td>
<td>-0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w12</td>
<td>1</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td>w41</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w13-16</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w42</td>
<td>-1</td>
<td>-0.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w17</td>
<td>0.9</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td>w43</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w18</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w44</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w19</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td></td>
<td>w45</td>
<td>-1</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w20</td>
<td>0.8</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td>w46-48</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w21</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td>w49</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w22</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td>w50</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w23-24</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w51</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w25</td>
<td>-0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w52</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>w26</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td></td>
<td>w53</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>w27</td>
<td>-0.2</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td>w54-55</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.4.2. Scenario 1: Satisfactory predicted action effect but intentionally inhibited action

The first scenario considered here describes a situation where the context is self, and a stimulus occurs. The simulation of this scenario is shown in Fig. 2. The predicted action effect of , is considered positive for the agent. Parallel to the objective prediction process, the subjective impact prediction process takes place and leads to an intentional inhibition.

In Fig. 2 it is shown that (after sensing the stimulus) the agent triggers preparation of action (in objective terms: ). Based on that the sensory representation of the predicted effect of is generated (through the as-if body loop), followed by activation of the feeling of , also in relation to the activated performative desires for . Next, these states contribute to generate activation of a prior self-ownership state (ownership of the action for the agent itself, attributed by the agent; this is in contrast with other-ownership: ownership of the action for another agent, attributed by the agent). In addition to the prior self-ownership state, a prior self-awareness state is developing, mainly upon the formation process of (objective) effect prediction and (subjective) impact prediction (Parkinson & Haggard, 2014). This development shows in the first (top) graph of Fig. 2 while the bottom graph shows the development of impact prediction.

With the activated constitutive desires for the agent activates preparation for abandoning the action, and this in turn has an increasing effect on the sensory representation of predicted effect of , followed by the feeling of over the feeling of which contributed to the developed very high prior awareness. Therefore, the preparation for abandoning the action is strengthened more and action execution is intentionally abandoned, as explained in (Brass & Haggard, 2007, 2008; Filevich, Kühn, & Haggard, 2012; Kühn, Haggard, & Brass, 2009; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012). This shows that there is no effect of an execution of action (via the body loop) in positive manner via the sensory representation of , and thus the feeling of with sensory representation of predicted effect remain at the same level while developing retrospective awareness. While maintaining the same level for sensory representation of predicted effect and feeling of it strengthen the

| Table 5. Steepness (σ) and threshold (τ) values used in configurations of simulations. |
|---------------------------------|---------|---------|---------|---------|---------|---------|---------|---------|
| PA(a_{obj})                    | PA(a_{sub}) | SR(b_{obj}) | SR(b_{sub}) | F(b_{obj}) | F(b_{sub}) | PD(b)   |
| σ                              | 1.5     | 1.2     | 3        | 3        | 3        | 3        | 4        |
| τ                              | 0.1     | 0.6     | 0.1      | 0.2      | 1        | 1        | 0.1      |
| CD(b)                          | PO      | PAwr    | EA(a)    | RO       | RAwr     | EO      |
| σ                              | 6       | 6       | 4        | 6        | 4        | 8        | 6        |
| τ                              | 0.3     | 1       | 0.9      | 1        | 1.8      | 1.35     | 0.7      |

3.4.2. Scenario 1: Satisfactory predicted action effect but intentionally inhibited action

The first scenario considered here describes a situation where the context is self, and a stimulus occurs. The simulation of this scenario is shown in Fig. 2. The predicted action effect of , is considered positive for the agent. Parallel to the objective prediction process, the subjective impact prediction process takes place and leads to an intentional inhibition.

In Fig. 2 it is shown that (after sensing the stimulus) the agent triggers preparation of action (in objective terms: ). Based on that the sensory representation of the predicted effect of is generated (through the as-if body loop), followed by activation of the feeling of , also in relation to the activated performative desires for . Next, these states contribute to generate activation of a prior self-ownership state (ownership of the action for the agent itself, attributed by the agent; this is in contrast with other-ownership: ownership of the action for another agent, attributed by the agent). In addition to the prior self-ownership state, a prior self-awareness state is developing, mainly upon the formation process of (objective) effect prediction and (subjective) impact prediction (Parkinson & Haggard, 2014). This development shows in the first (top) graph of Fig. 2 while the bottom graph shows the development of impact prediction.

With the activated constitutive desires for the agent activates preparation for abandoning the action, and this in turn has an increasing effect on the sensory representation of predicted effect of , followed by the feeling of over the feeling of which contributed to the developed very high prior awareness. Therefore, the preparation for abandoning the action is strengthened more and action execution is intentionally abandoned, as explained in (Brass & Haggard, 2007, 2008; Filevich, Kühn, & Haggard, 2012; Kühn, Haggard, & Brass, 2009; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012). This shows that there is no effect of an execution of action (via the body loop) in positive manner via the sensory representation of , and thus the feeling of with sensory representation of predicted effect remain at the same level while developing retrospective awareness. While maintaining the same level for sensory representation of predicted effect and feeling of it strengthen the

| Table 5. Steepness (σ) and threshold (τ) values used in configurations of simulations. |
|---------------------------------|---------|---------|---------|---------|---------|---------|---------|
| PA(a_{obj})                    | PA(a_{sub}) | SR(b_{obj}) | SR(b_{sub}) | F(b_{obj}) | F(b_{sub}) | PD(b)   |
| σ                              | 1.5     | 1.2     | 3        | 3        | 3        | 3        | 4        |
| τ                              | 0.1     | 0.6     | 0.1      | 0.2      | 1        | 1        | 0.1      |
| CD(b)                          | PO      | PAwr    | EA(a)    | RO       | RAwr     | EO      |
| σ                              | 6       | 6       | 4        | 6        | 4        | 8        | 6        |
| τ                              | 0.3     | 1       | 0.9      | 1        | 1.8      | 1.35     | 0.7      |

3.4.2. Scenario 1: Satisfactory predicted action effect but intentionally inhibited action

The first scenario considered here describes a situation where the context is self, and a stimulus occurs. The simulation of this scenario is shown in Fig. 2. The predicted action effect of , is considered positive for the agent. Parallel to the objective prediction process, the subjective impact prediction process takes place and leads to an intentional inhibition.

In Fig. 2 it is shown that (after sensing the stimulus) the agent triggers preparation of action (in objective terms: ). Based on that the sensory representation of the predicted effect of is generated (through the as-if body loop), followed by activation of the feeling of , also in relation to the activated performative desires for . Next, these states contribute to generate activation of a prior self-ownership state (ownership of the action for the agent itself, attributed by the agent; this is in contrast with other-ownership: ownership of the action for another agent, attributed by the agent). In addition to the prior self-ownership state, a prior self-awareness state is developing, mainly upon the formation process of (objective) effect prediction and (subjective) impact prediction (Parkinson & Haggard, 2014). This development shows in the first (top) graph of Fig. 2 while the bottom graph shows the development of impact prediction.

With the activated constitutive desires for the agent activates preparation for abandoning the action, and this in turn has an increasing effect on the sensory representation of predicted effect of , followed by the feeling of over the feeling of which contributed to the developed very high prior awareness. Therefore, the preparation for abandoning the action is strengthened more and action execution is intentionally abandoned, as explained in (Brass & Haggard, 2007, 2008; Filevich, Kühn, & Haggard, 2012; Kühn, Haggard, & Brass, 2009; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012). This shows that there is no effect of an execution of action (via the body loop) in positive manner via the sensory representation of , and thus the feeling of with sensory representation of predicted effect remain at the same level while developing retrospective awareness. While maintaining the same level for sensory representation of predicted effect and feeling of it strengthen the
idea that although the selected action was intentionally abandoned, the factors that contributed to the emergence of that selection will not disappear suddenly (Filevich, Kühn, & Haggard, 2012). Further it is observed that no retrospective ownership developed. Finally, the agent communicates self-ownership about the abandoned action based on retrospective self-awareness. Note that when the stimulus is taken away, all activation levels will come down to 0 (q.v. LP1), and will come up again when the stimulus reoccurs. Aligning with the observation by Walsh et al. (Walsh, Kühn, Brass, Wenke, & Haggard, 2010); when intentional inhibition occurred an additional time has been consumed: this can be observed by comparing the timelines of this with Scenario 2.

Fig. 2: Scenario 1: Satisfactory predicted action got intentionally inhibited. Both top and bottom graphs are on the same simulation and have been presented separately to improve the readability and clarity.
3.4.3. Scenario 2: Action with satisfactory predicted effect gets executed

The second scenario is identical to the first but intentional inhibition does not occur and the action will be performed. This scenario describes a situation where the context \( c \) is the agent itself, and a stimulus \( s \) occurs. The action effect \( b \) of \( a \), is considered positive for the agent. Parallel to this the impact prediction process takes place and evaluates the appropriateness on the action selection from long term perspectives (given the constitutive desires). The simulation of this scenario is shown in Fig. 3.

In Fig. 3 it is shown that (after sensing the stimulus) the agent triggers preparation of action \( a_{obj} \). Based on that the sensory representation of the predicted effect \( b_{obj} \) of \( a \) is generated (through the as-if body loop) and followed by the feeling \( F(b_{obj}) \) of \( b_{obj} \) with the aid of the activated performative desire for \( b \). Next, these

![Fig. 3: Scenario 2: Satisfactory predicted action got executed while impact prediction is enabled. Both top and bottom graphs are on same simulation and have been presented separately to improve the readability and clarity.](image)
Chapter 3

states contribute to generate a prior self-ownership. With the prior self-ownership, prior self-awareness is also developing, mainly upon the formation process of (objective) effect prediction and (subjective) impact prediction. This development shows in the first (top) graph of the Fig. 3 while the bottom graph of the same figure shows the development of impact prediction. With the activated constitutive desires for b agent strengthens preparation PA(a_{sub}) for abandoning action a. However, in the evaluation process through the impact prediction link it turns out the impact of action a is valued as not negative (the values of SR(b_{sub}) and F(b_{sub}) stay low). This low level not contribute positively to the preparation PA(a_{sub}) for abandoning action a. This preparation has not gone to the activation level high enough to enable intentional inhibition. Therefore, agent performs the actual execution of action a which propagates its effects through the body loop (P. Haggard & Eimer, 1999; Banks & Isham, 2009). In Fig. 3 it clearly shows that the execution of action a (via the body loop) also affects in positive manner via the sensory representation SR(b_{obj}) and the feeling F(b_{obj}) of b_{obj}: the sensory representation of b gets further strengthen from action execution. In parallel the sensory representation of b_{obj} is suppressed due to the prior self-ownership state which causes a slight relative dip in the graph of the sensory representation of b_{obj} between time points 25 and 33 (Blakemore, Wolpert, & Frith, 2000; Fourneret et al., 2002). Due to the action execution the agent develops a retrospective self-ownership state which is followed by a retrospective self-awareness state. Finally, the agent communicates self-ownership about the performed action based on retrospective self-awareness and ownership.

3.4.4. Scenario 3: Action with satisfactory predicted effect gets executed, while impact prediction is impaired

The third scenario shows the backward compatibility of the model with (Thilakarathne & Treur, 2013b) which does not have the impact prediction. Therefore, to simulate this, the connections related to impact prediction were disabled as per the relevant connection weight changes shown in Table 4 under the column S3. The scenario describes a situation where the context c is the agent itself, and a stimulus s occurs. The action effect b (obj) of a, is considered positive for the agent and the awareness of action formation and execution will be scrutinized together with generated prior and retrospective ownership states. The simulation of this scenario is shown in Fig. 4.

In Fig. 4 it is shown that (after sensing the stimulus) the agent triggers preparation of action a with the aid of triggered performative desires for b. Based on that the sensory representation of predicted effect b (obj) of a is generated (through the as-if body loop) and followed by the feeling of b. Next these states contribute to generate a prior self-ownership. After activating the prior self-ownership, prior self-
awareness is developing, mainly upon the formation process of effect prediction $b$ of $a$. After that, as a result of prior self-awareness and ownership states, the agent initiates the actual execution of action $a$ which propagates its effects through the body loop (P. Haggard & Eimer, 1999; Banks & Isham, 2009). In the Fig. 4 it clearly shows that the execution of action $a$ (via the body loop) also affects in positive manner via the sensory representation $b$ (obj) of $a$ and the feeling of $b$ (obj) (sensory representation of $b$ get further strengthen from action execution). In parallel the sensory representation $b$ (obj) of $a$ is suppressed due to the prior self-ownership state which causes a slight dip in the graph (Blakemore, Wolpert, & Frith, 2000; Fourneret et al., 2002)]. Due to the action execution the agent develops a retrospective self-ownership state which is followed by a retrospective self-awareness state. Finally, the agent communicates self-ownership about the performed action based on retrospective self-awareness and ownership. From the

![Diagram](image.png)

**Fig. 4**: Scenario 3: Satisfactory predicted action gets executed while impact prediction is impaired. Both top and bottom graphs are on same simulation and have been presented separately to improve the readability and clarity.
Fig. 4 bottom graph it clearly shows that all the states attached to the impact prediction loop do not have any activation through that time period.

3.4.5. Scenario 4: Model behaviour in total unconscious mode

The fourth scenario shows the total unconscious mode of the model (i.e., no impact prediction process and awareness states become active). The parameter values changed to have this mode are with the Table 4 under S4 column. In here also the scenario describes a situation where the context \( c \) is the agent itself, and a stimulus \( s \) occurs. The action effect \( b(\text{obj}) \) of \( a \), is considered positive for the agent. The simulation of this scenario is shown in Fig. 5.

In Fig. 5 it is shown that (after sensing the stimulus) the agent triggers preparation of action \( a \) with the aid of triggered performative desires for \( b \). Based on that the sensory representation of predicted effect \( b(\text{obj}) \) of \( a \) is generated (through the as-if body loop) and followed by the feeling of \( b \). Next these states contribute to generate a prior self-ownership. After that, the agent initiates the actual execution of action \( a \) which propagates its effects through the body loop (Haggard & Eimer, 1999; Banks & Isham, 2009). In the Fig. 5 it clearly shows that the execution of action \( a \) (via the body loop) also affects in positive manner but with a relatively low degree of strength (relative to the scenario two and three) via the sensory representation \( b(\text{obj}) \) of \( a \) and the feeling of \( b(\text{obj}) \). This gives the idea that when performing an action with the awareness the strength of the action is high (if it is positively formed). In parallel the sensory representation \( b(\text{obj}) \) of \( a \) is suppressed due to the prior self-ownership state which causes a slight dip in the graph [cf. (Blakemore, Wolpert, & Frith, 2000; Fourneret et al., 2002)]. Due to the action execution the agent develops a retrospective self-ownership state. Fig 5 clearly shows that all states attached to the impact prediction loop and awareness states do not have any activation.

![Fig. 5: Scenario 4: Model behavior in total unconscious mode.](image)
3.4.6. Scenario 5: Inability to proactively use constitutive desires

This fifth scenario shows the importance of constitutive desires, together with subjective action preparation PA($a_{sub}$) for the intentional inhibition process. In principle, if the agent is unable to deploy the processes addressing the subjective aspects (the impact prediction and evaluation processes) with adequate strength, it will not be able to properly inhibit the action which is going to be executed. This simulation uses the same configuration as in the first scenario but only with two changes related to constitutive desires CD($b$) and subjective action preparation PA($a_{sub}$). These weight changes are shown in Table 4 under column S5. In this scenario also the context $c$ is the agent itself, and a stimulus $s$ occurs. The action effect $b$ of $a$ is considered positive for the agent. Parallel to this, the impact prediction process takes place and evaluates the appropriateness of the action.

**Fig. 6**: Scenario 5: Inability to proactively use constitutive desires. Both top and bottom graphs are on same simulation and have been presented separately to improve the readability and clarity.
selection from long term perspectives, but with a impairment concerning constitutive desires. The simulation of this scenario is shown in Fig. 6.

In Fig. 6 it is shown that (after sensing the stimulus) the agent triggers preparation of action $a_{obj}$. Based on that, the sensory representation of the predicted effect $b_{obj}$ of $a_{obj}$ is generated (through the as-if body loop) and followed by the feeling $F(b_{obj})$ of $b_{obj}$ with aid of the activated performative desire for $b$. Next, these states contribute to generate a prior self-ownership state. With this prior self-ownership state, prior self-awareness is also developing, mainly upon the formation process of (objective) effect prediction and (subjective) impact prediction. This development of effect prediction shows in the first (top) graph of Fig. 6 while the bottom graph of the same figure shows the development of impact prediction. The poor activation of constitutive desires for $b$ has further effects on the strength of preparation $PA(a_{sub})$ for abandoning action $a$. However, in the valuation process through the impact prediction link the impact of action $a$ is valued as low: the values of $F(b_{sub})$ stay very low. This low level does not contribute positively to the preparation $PA(a_{sub})$ for abandoning action $a$. This preparation has not gone to an activation level high enough to enable intentional inhibition properly in this scenario. Therefore, the action has successfully been executed, including the required retrospective states. As this scenario is identical to the first scenario (except for two weights), this shows how important constitutive desires are for the intentional inhibition process.

3.5. Discussion

Having some form of control over your actions is crucial in many areas of life. These areas vary from functioning adequately in a social context to performing in an effective manner at work, maintaining a healthy lifestyle, and maintaining a sustainable lifestyle. In such cases poor control or complete lack of control can lead to, for example, socially unaccepted actions, spending too much time at work on work-irrelevant actions, actions that damage health such as taking unhealthy food or drugs, or actions that damage our global life environment: the earth.

Our environment usually presents us many stimuli that trigger undesirable actions. Our cognitive system responds to these stimuli by initiating preparations for such actions. However, our internal control mechanisms – if functioning properly - may refrain us from actually executing such prepared actions. In recent years much detailed knowledge on these mechanisms has been developed, especially in the cognitive neuroscience area. However, existing cognitive computational models often use less detailed knowledge from earlier years, without taking into account the rich source of knowledge given by the biological perspective. The current paper shows how the more detailed recently developed knowledge can be used to obtain a computational model.
The computational model presented here was inspired both by cognitive and neurological evidence, and has shown the combined impact from intentional inhibition, action awareness, and action ownership. The intentional inhibition process provides a core capability to demonstrate self-control in a situational context that confirms a surviving ‘social entity’ (Mostofsky & Simmonds, 2008; Cohen, Berkman, & Lieberman, 2013). In parallel to the positive action selection process, the intentional inhibition process evaluates the possible negative influence of the current action selection from the long term perspective and may lead to abandoning the action. Therefore, intentional inhibition leads to stopping one’s own action without any external clues or signals and this is defined as a capacity to voluntarily suspend or inhibit an action (Filevich, Kühn, & Haggard, 2012). This is completely different from classical psychological process behind No-Go tasks (Eimer, 1993; Pfefferbaum, Ford, Weller, & Kopell, 1985). This paper includes supporting states to bring about the effects of intentional inhibition. Among those supporting states performative desires (which represent short term interests/goals) and constitutive desires (which drives a person through his or her long term driven aspirations) play a key role to provide the above mentioned positive and negative evaluations. Furthermore, ownership both from a prior and retrospective perspective contributes to the action formation process from an unconscious perspective. The awareness from both a prior and retrospective contributes to intentional inhibition from a conscious perspective. This interplay between conscious and unconscious processes is emphasized, for example, in (Dehaene & Naccache, 2001; Patrick Haggard, 2008; Kiefer et al., 2011; Winkielman & Schooler, 2011; Cleeremans, 2011).

To validate the model behaviour, five scenarios have been simulated. The model was translated into a computational dynamical system, and a generic parameter value set was identified through a systematic analytical approach, in order to demonstrate the behaviour of each scenario. By having such a generic, global parameter value set the strength of current model in practical application contexts is shown, and how strongly it correlates with the theoretical basis that was used to design this model. The simulation results are in line with the literature discussed in Section 1 and 2. Scenario 1 presents a situation in which an action with satisfactory predicted action effect is intentionally inhibited. This is the core functionality of this model; once the agent has found satisfactory predicted effects of an action (i.e., agent is having a strong positive feeling of the effects of executing the selected option before actually executing it), this shows the ability of inhibiting an action which is going to be executed. Furthermore, this scenario highlights the interplay between effect prediction and impact prediction as a cognitive process. The agent has used performative desires to prepare for the action through the effect prediction loop, while it is evaluating the impact of the same action through constitutive
Chapter 3

desires. When there is a mismatch or conflict between these two parallel processes, this will lead to inhibition of the action. Scenario 2 is the same as the scenario 1 but the agent is executing the action this time. This shows that if there is no mismatch (or the difference is not significant) in the effect prediction and impact prediction processes, not inhibition will take place. In this simulation the agent has developed a strong predictive effect on the selected action and it has successfully executed the same action with adequate retrospective effects, in line with its prior effects. Scenario 3 is providing information of the influence of impact prediction on the intentional inhibition process. In this scenario the impact prediction loop is not working and the configuration of scenario 1 is executed. In scenario 1, having both effect prediction and impact prediction it is shown that the agent inhibits the action, nevertheless; under the same configuration but by absence of the impact prediction, scenario 2 shows that the agent is executing the selected action. This helps to show the effects of impact prediction on action inhibition and confirms the interplay between effect prediction and impact prediction and the importance of both to have adequate intentional inhibition. Scenario 4 provides information on the strength of conscious and unconscious influences on the proposed model. In all the previous scenarios the agent developed awareness on the selected action and later decides to inhibit or not to inhibit the selected action. Nevertheless, in complex simulations it is necessary to have the effects of unconscious and conscious effects of action selection. While having three scenarios showing the influence of awareness state, the 4th scenario provides a situation in which the agent has no impact prediction process nor awareness states. Therefore, this scenario provides the behaviour of the model in total unconscious mode. The results provide realistic outcomes; the agent executes the action but with relatively lower strength (having no awareness, the selection process leads to automatic action selection where the strength may be a bit weaker in most of the cases). The final scenario shows the importance of constitutive desires specifically. In this situation the impact prediction loop works well, but the agent is unable to bind a strong constitutive desire (this can be considered as a mental impairment situation). Having an identical configuration as in scenario 1, due to poor activation strength for constitutive desires (with the changes highlighted in Table 4) the agent is unable to deploy adequate strength of the subjective aspects, which leads to not being able to properly inhibit the action which is going to be executed. This shows the key role of the constitutive desires on the effect prediction loop; if a person is not having strong long term desires it is more difficult to cope with a situation compared to a healthy person.

The experiments have highlighted the fact that if intentionally an action was abandoned it takes relatively more time to get settled with the original feelings compared to the same when the action got successfully executed (see Scenario 1 and 2). The Anarchic Hand Syndrome (AHS) is a neurological disorder (Filevich,
Kühn, & Haggard, 2012) that can be simulated analogically by considering Scenario 1 and 2 also. This type of more complex simulations are future extensions for this model and specially by incorporating emotion-related cognitive processes.

The agent model is meant as a basis for subsequent work on developing ambient agent systems able to monitor, analyse and support persons trying to develop a healthy lifestyle. If such systems have such a model of the underlying human processes, they can use this to have a more deep understanding of the human. As possible application domains the following will be addressed: decision making, behavioral management, emotional control, and simulations for clinical disorders and therapies for them. Furthermore, this model includes detailed information on the intentional inhibition process, within the action selection processes. Therefore, the complexity of this model is high though it has many advantages when comparing it to highly abstract and simple cognitive models. The selection of parameter values for this model is a non-trivial task and due to the high order of coupling among states, just selecting arbitrary values will not provide expected results. Therefore a systematic analytical approach had to be and actually was used for this purpose. However, it is difficult to directly validate each of these selected values for parameters based on empirical data about it. Having a generic value set for parameters and providing five distinct scenarios where each is interrelated (i.e., relative to the first scenario, other scenarios are different only by a few changes in specific weigh values in connection links for states depending on the focus of that particular scenario) this choice can be justifiable. This shows the value of the model in applications: it is easy to explore different situations just with a few changes to the parameter values relative to the reference scenario.

There is also interesting literature evidence available on emotional influence on action inhibition and selection for better self-sustainability and non-social dysfunction; e.g., (Kimhy et al., 2012). In future research it is planned to extend this model by adding the cognitive and neurological evidence on explicit and implicit emotional regulation, with affective, and social consequences; cf. (Lynn, Muhle-Karbe, & Brass, 2014).

References
http://doi.org/10.1523/JNEUROSCI.4682-05.2006


Jahfari, S., Waldorp, L., van den Wildenberg, W. P. M., Scholte, H. S., Ridderinkhof, K. R., & Forstmann, B. U. (2011). Effective Connectivity Reveals Important Roles for Both the Hyperdirect (Fronto-Subthalamalic) and the Indirect (Fronto-Striatal-Pallidal)
http://doi.org/10.1523/JNEUROSCI.5253-10.2011

http://doi.org/10.2478/v10053-008-0090-4

http://doi.org/10.1016/j.psychres.2012.05.029

http://doi.org/10.1016/j.cogbrainres.2005.03.015

http://doi.org/10.1016/j.neuroimage.2009.03.020

http://doi.org/10.1002/hbm.20711

http://doi.org/10.1126/science.1090973

http://doi.org/10.1093/brain/106.3.623

http://doi.org/10.1016/j.neuropsychologia.2014.09.009

http://doi.org/10.1016/j.concog.2006.12.004

http://doi.org/10.1162/jocn.2008.20500

http://doi.org/10.1037/0022-3514.91.3.524

http://doi.org/10.1016/j.cognition.2007.09.003


Chapter 4

Modelling the Dynamics of Emotional Awareness\(^1\)

Dilhan J. Thilakarathne, Jan Treur

Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.
e-mail: d.j.thilakarathne@vu.nl, j.treur@vu.nl

**Abstract:** In this paper, based on literature from Cognitive and Affective Neuroscience, a computational agent model is introduced incorporating the role of emotional awareness states in the dynamics of action generation. More specifically, it covers both automatic, unconscious (bottom-up) and more cognitive and conscious (top-down) emotion generation processes, and their mutual interaction. The model was formalised in a dynamical system format. In different scenarios the model shows simulation results that are in line with patterns reported in literature.

---

\(^1\) This chapter was published as:
http://doi.org/10.3233/978-1-61499-419-0-885

*The names of the authors are ordered alphabetically reflecting the comparable contribution of each author.*
4.1. Introduction

Generation of emotions may take place by automatic processes (unconscious, referred to as bottom-up) and/or by conscious processes (with awareness, referred to as top-down) [11, 16]. The relation between emotion and conscious awareness is a nontrivial one. Latest findings suggest that conscious influences of emotion are playing a role that should not be underestimated (cf. [14]). To address this a Levels of Emotional Awareness Scale (LEAS) to quantify emotional experience together with to elaborate emotional experience of individuals was introduced (cf. [14]). Lack of emotional awareness is considered a main factor behind many emotional disorders (e.g., alexithymia [25], schizophrenia [3]), and having insight in the neurological and behavioural basis of emotional awareness will support the understanding of the process behind this innate ability of living beings [13]. In this paper a neurologically and behaviourally inspired computational model is introduced together with neural correlates as a set of affective states that is able to describe and simulate the dynamics of emotional awareness in interaction with perception, attention, and preparing and performing actions. The precise functional contribution of the neural regions indicated in this paper may need further research and confirmation. Nevertheless, the discussed body of knowledge might be useful as a basis for a workbench for the AI community to strengthen some of the intelligent applications addressing human-like processes, and also to provide an experimental framework for neuro-cognitive-behavioral scientists.

4.2. Neurological Background

Emotion formation is an ongoing process and not necessarily is triggered in merely an instant [10]. Emergence of emotions has different explanations, based on its automatic responses (bottom-up), or more consciously emerging (top-down). These approaches have been able to explain emotional formation in line with results from fMRI experiments [11, 16]. Examples for this bottom-up and top-down mechanisms (from [23]) are experiencing disgust as a result of smelling outdated milk and recollecting smelling outdated milk, respectively. Bottom-up emotion generation is assumed to be aroused immediately and ingrained from an external stimulus while the top-down emotion generation occurs from semantic evaluation of a situation through a cognitive influence [23]. It has been shown that different neural activations are evoked for this: thalamus, hypothalamus, ventral striatum, amygdala, anterior cingulate cortex (ACC), anterior insular cortex (AIC), orbitofrontal cortex (OFC), and/or mesial prefrontal cortex [14], with and without conscious intervention.

Evidence was found for the idea of distributed networks of regions collectively carrying out important functions of the brain (no single regions), including emotion
generation [2, 31]. The amygdala is the main hub not only for monitoring the emotionally salient stimuli but also for projecting to the relevant brain areas (it has connectivity with eight of the cortical areas [17]) and transmit retrieved feedbacks to the sensory pathways, to invoke rapid and efficient generation of emotions [5, 18, 21]. The amygdala may have an important contribution when processing danger (e.g., flight or fight situation) or emotionally salient events, especially when these occur outside attention or awareness [31]. From available fMRI data it is noted that the left amygdala seems to directly contribute for both bottom-up and top-down processes and the right amygdala has shown activation only for the bottom-up responses [16]:

‘.. distinct cortical networks were involved in each type of emotion generation. On the one hand, bottom-up emotion generation activated the amygdala and occipital cortex, which have been implicated in detecting affectively arousing stimuli and modulating their encoding into memory (..), as well as right prefrontal and parietal regions implicated in attentional vigilance and individual differences in negative affective style (..). On the other hand, top-down emotion generation activated left prefrontal, cingulate, and temporal regions implicated in working memory and the retrieval of information from semantic memory (..), as well as the left amygdala and a dorsal mPFC region involved in making attributions about mental—and especially emotional—states (..). Working together, these systems may support cognitive appraisals that generate emotions from the top down.’ ([16], pp. 1327-1328).

According to [18] the amygdala directly shapes the perception when perceiving an emotionally salient stimulus (bottom-up) and from [20] emotional perception contributes to identify emotionally salient information in the environment, and to generate emotional experiences and behaviour, and also from [21] emotions can be shaped by the perception through amplification mechanisms that do not overlap with other attentional processes (without leading to awareness). In this bottom-up process the brain has shown to capture the emotional perceptual features of the stimulus spontaneously but not involving conscious awareness and further subjective aspects of this emotion [15]. Therefore, perception may compel to emotion generation in the bottom-up approach.

Feelings which concern subjective experience of emotions [7] also play their role in different ways. The insula and ACC are believed to be neural correlates of feelings [7]:

‘While emotions are actions accompanied by ideas and certain modes of thinking, emotional feelings are mostly perceptions of what our bodies do during the emoting, along with perceptions of our state of mind during that same period of time.’ ([7], pp. 110).
It may possible to have different feelings on a perceived stimulus due to pre-learned neural paths (cf. [7, 8]); only a few of them may be able to reach consciousness through attention [7, 21].

Attention is a key cognitive process that allows (by subjectively desiring) to appraise a situation with conscious awareness [26]. While perception is a key aspect in the bottom-up process, attention compels the top-down process. As information processing through perceptual pathways is to be limited, attention contributes to select the most useful information and let it reach conscious awareness [21] (these types of emotions have shown a higher scores in LEAS [14]). Furthermore, there are mainly two types of attention mechanisms: exogenous (for bottom-up) and endogenous (for top-down); with partly distinct brain circuits [21, 32]. For attention also a networked brain region has been suggested involving frontoparietal regions (see [17, 6, 32]).

It has been noted that people with a high level of emotional awareness have shown to accurately detect and discriminate emotional signals [14], and [7] has shown the advantage of conscious awareness of emotion when integrating in cognitive processes; [12] has presented four evidences for emotional awareness and its conscious experience:

‘1) AIC and ACC are commonly coactivated as revealed by a meta-analysis, 2) AIC is functionally dissociable from ACC, 3) AIC integrates stimulus-driven and top-down information, and 4) AIC is necessary for emotional awareness.’ ([12], p. 3371).

Also it has been identified that the right-AIC, ventromedial Pre-Frontal Cortex (vmPFC), and ACC play a role as shared neural substrates for the awareness of bodily and emotional states (see [28]). Furthermore, for the bottom-up responses they found activity in the right-PFC (may relate to attention shifting) whereas for the top-down processes activity in dorsal left-PFC is observed [16] (may relate to semantic processing with awareness [22]). The importance of improving the emotional awareness in clinical perspectives has been highlighted, for schizophrenia [3, 13], alexithymia [25], and other cognitive disorders.

The OFC and Cholinergic Nuclei are noted to contribute to boosting emotional perceptual processing; and amygdala, fusiform gyrus, dorsolateral-PFC and inferior parietal cortices are for emotional awareness (cf. [1, 5]). It has clearly been shown that emotional perception is modulated by attention [19]; the typical interplay between attention and consciousness and/or awareness can be found in [9, 27].

4.3. Description of the Model

With the evidence presented in Section 2, a computational agent model has been designed for emotion formation which adopts parts of the models presented in [29, 30] but extends those by introducing emotional awareness in interaction with
perception, and attention. An overview of this model is shown in Figure 1. Modeling causal relations discussed in neurological literature in the manner as presented here does not take specific neurons/paths into consideration but uses more abstract cognitive or mental states. The model uses three world states as inputs: for stimulus \( s \), context \( c \), and effect \( b \). These inputs; world states WS\((s)\), WS\((c)\), and WS\((b)\) lead to sensor states SS\((s)\), SS\((c)\), and SS\((b)\), and subsequently to sensory representation states SR\((s)\), SR\((c)\), and SR\((b)\), respectively. This initiation propagates through two causal chains as proposed by Damasio [8] (for more details see [29, 30]):

- as-if body loop [preparation for action \( a \): PA\((a) \rightarrow SR(b) \rightarrow \text{feeling of action } a \text{ after: as-if loop or body loop: } F(b)\)]
- body loop [PA\((a) \rightarrow \text{execution of action } a \): EA\((a) \rightarrow WS(b) \rightarrow SS(b) \rightarrow SR(b) \rightarrow F(b)\)]

The effect prediction as-if body loop contributes to action selection in a parallel mode, i.e., developing preparations PA\((a_i)\) for a number of actions \( a_i \) where \( i = 1, \ldots, n \). These multiple action candidates \( a_i \) are competing to get selected [7, 8]. Furthermore, this model takes the influence from performative desires for \( b \): PD\((b)\) on PA\((a)\) and F\((b)\) to introduce the influence from short term interests/goals for selecting or rejecting an action through the as-if body loop (cf. [29]). As in [29, 30] these loops have been extended with prior and retrospective effects relative to the

![Fig. 1: Overview of the computational cognitive agent model. Red colour ➔ and ─ symbols presenting suppressions; and Y:- a,b,c,e,s.](image-url)
action execution through ownership and awareness states (this model includes emotional awareness also into this). The prior or retrospective-ownership state for action \( a \) with \( b, c, e, \) and \( s \): \( \{\text{PO} \} O(Y= a,b,c,e,s) \) was to represent in how far a person attributes an action to him/herself, or to another person, whereas the prior or retrospective-awareness state for action \( a \) with \( b, c, e, \) and \( s \): \( \{\text{Awr} \} Awr(Y= a,b,c,e,s) \) for the influence of conscious elements (cf. [29, 30]). Apart from the relations presented in [29, 30] this model covers two causal emotion formation processes: bottom-up and top-down. Emotional perception [5, 9, 18, 21], and attention [19, 21, 26, 27, 32] with emotional awareness [9, 12, 22, 27, 28] have been found as key factors contributing to these emotion formation processes.

4.3.1. Bottom-up process

In the bottom-up process, when a particular stimulus \( s \), and a context \( c \) (which are emotionally salient) are perceived, the agent will spontaneously develop an emotional perception state for \( s \) with \( b \): \( \text{EPer}(s,b) \) together with an influence from performative desires \( \text{PD}(b) \) [15, 21]. Subsequently preparation \( \text{PA}(e) \) for emotional response \( e \) and preparation state \( \text{PA}(a) \) are independently affected by the perception state \( \text{EPer}(s,b) \) [18, 20, 21]. Furthermore, the preparation state \( \text{PA}(a) \) is affected by \( \text{PA}(e) \) too, so that if there is a strong perception that directly strengthens the action preparation [17] leading to a spontaneous response with a higher strength (e.g., flight or fight [24]). This preparation state \( \text{PA}(a) \) triggers the effect prediction sub-process (as-if body loop) that internally generates a sensory representation of the bodily response and feeling for the associated emotions before actually executing the action [8].

Based on the internally simulated feeling state \( \text{F}(b) \) an emotional attention state \( \text{EAtt}(b) \) will be developed for the current selection of action \( a \) and its effect \( b \) [18, 21]. Nevertheless, the state \( \text{EAtt}(b) \) is not a main factor affecting \( \text{EPer}(s,b) \) in the bottom-up process [21]. The preparation state \( \text{PA}(e) \) is affected by the feeling state \( \text{F}(b) \) and therefore this will contribute to select the action due to satisfactory valuation together with the direction of the perception state [21]. In the bottom-up process, the strong activation level of \( \text{EPer}(s,b) \) is developed in an early stage of the timeline and therefore dominates the mechanism; see also, e.g., [9, 15, 21]. Due to the necessity of immediate and strong action execution in the bottom-up process the cognitive appraisal sub process on action selection (through the as-if body loop) may not significantly contribute whereas the strong activation level of \( \text{EPer}(s,b) \) will direct to an action preparation without getting biased from emotional attention. In this process the brain is directed to rationally engage in the big picture of the current threat [15].

A prior ownership state \( \text{PO}(a,b,c,e,s) \) is affected by \( \text{SR}(c) \) (see [30]), \( \text{PA}(a) \), \( \text{PA}(e) \), and \( \text{F}(b) \). The ownership state is contributing to the unconscious aspects (see
and especially in the bottom-up process this explains the aspects of negative emotions (e.g., fear). For example, in a flight or fight situation though the agent knows that it is performing an action (through the ownership), he/she may not be really sure why it is doing so (due to lack of awareness). In the meantime the agent will perform emotional expression EE(e) of e, as a result of PA(e). Furthermore, based on the prior ownership state PO(a,b,c,e,s) and preparation PA(a) the agent will perform execution EA(a) of action a through the body loop. The sensory representation state SR(b) will be suppressed once the prior ownership state PO(a,b,c,e,s) got developed (as explained in [29, 30]); this allows to differentiate effects on SR(b) from the as-if body loop against effects from the body loop (see [30]). Subsequently a retrospective ownership state RO(a,b,c,e,s) for action a with b, c, e, and s will be developed; this is affected by PO(a,b,c,e,s), F(b), and EA(a). Both EPer(s,b) and PO(a,b,c,e,s) are suppressed by the effects of RO(a,b,c,e,s); therefore agent will able to dilute the strength of action (after the execution) from retrospective effects. Communication EO(a,b,c,e,s) of ownership for action a with b, c, e, and s is affected only by the RO(a,b,c,e,s) state in the bottom-up process and agent will able to share the information with external agents. In this bottom-up process the agent will not experience any awareness state (PAwr or RAwr) for the emotion and/or action as in [9, 15, 21, 31].

4.3.2. Top-Down Process

With the influence from the world through a stimulus s, and a context c the agent will be prepared by PA(a) for an action a, in relation to performative desire PD(b) (cf. [29]). As an effect from PA(a), by internal simulation the agent will develop SR(b) and F(b) through the as-if body loop as suggested by Damasio in [7, 8]. The top-down process involves a role of subjective desires SD(b) [14, 15, 28], an early stage of the emotional attention state EAtt(b) development [6, 32] relative to the emotional perception state EPer(s,b) and the awareness states (emotional and action). Therefore, in parallel to the above action formation process, the agent is experiencing a salient activation of subjective desires SD(b) as an effect from the both SR(c) and SR(s). Subsequently the agent starts to develop an emotional attention state EAtt(b) for bi, giving attention to a particular bi [6, 32]. This particular bi may be a weak action candidate in the pool of parallel internal as-if body loop simulations. The term ‘appraisal’ in the literature occurs in this model through this valuation of parallel action simulations. Nevertheless, due to high attention developed for that bi it may strengthen more and more and beat all the other candidates [14, 15, 16] (modify or suppress these evaluations [31]). The emotional attention state EAtt(b) has an effect from the subjective desire SD(b) and vice versa [26]. Because of this emotional attention state EAtt(b), the agent will start to develop an emotional perception state EPer(s,b) (the perception of emotion
laden items requires attention, see [17]), and this leads to a preparation of an emotional response PA(e) too. Besides, PA(e) is affected by the feeling state F(b). Subsequently, the prior emotional awareness state PEA(b,e) of b and e develops due to the effects of SD(b), EAtt(b), PA(e) and F(b) which is another key state in the top-down process [12, 14]. As another consequence of the preparation state PA(e) the agent will develop an expressed emotional response EE(e) and experience the subsequent effects in terms of the feeling of it through the body loop [7, 8].

Together with the development of the prior emotional awareness state PEA(b,e), also the prior ownership state PO(a,b,c,e,s) will be developed as an effect from the states PA(e), SR(c), PA(a), and F(b) (this contributes to an interplay between conscious and unconscious processes in this model). Prior awareness PAwr(a,b,c,e,s) of a with b, c, e, and s is affected by the feeling state F(b) and the prior ownership state PO(a,b,c,e,s) (cf. [29, 30]). Subsequently, execution of the action a will be triggered as an effect of the states PO(a,b,c,e,s) and the PA(a). The retrospective emotional awareness REA(b,e) is affected by PEA(b,e), EE(e), F(b), and RAwr(a,b,c,e,s), and once this state REA(b,e) has developed it suppresses the emotional perception EPer(s,b) and subjective desire SD(b) to dilute the effects of current action formation. In parallel to that retrospective ownership RO(a,b,c,e,s) is affected by the states PO(a,b,c,e,s), F(b), and EA(a) (as in [29, 30]). Furthermore, the retrospective awareness state RAwr(a,b,c,e,s) is affected by the states REA(b,e), F(b), PAwr(a,b,c,e,s), RO(a,b,c,e,s), and EA(a). The prior awareness state PAwr(a,b,c,e,s) is suppressed by the retrospective awareness state RAwr(a,b,c,e,s). Finally, the communication (in retrospect) of ownership EO(a,b,c,e,s) is developed as an effect of the retrospective states RAwr(a,b,c,e,s), REA(b,e), and RO(a,b,c,e,s). These processes refer to the elicitation of emotions largely by cognitions through subjectively driven appraisal processes which are not primarily tied to a particular perceptual stimulus [15].

4.3.3. Dynamics of the model

Connections between the different state properties (the arrows in Figure 1) have weights $\omega_k$, as indicated in Table 1. In this table the column LP refers to the (temporally) Local Properties (LP) in LEADSTO format listed in the Extended Appendix\(^2\) (see [4] for the relevance and benefits of LEADSTO in dynamic models). A weight $\omega_k$ has a value between -1 and +1 and may depend on the specific context c, stimulus s, action a, effect b, and emotion e involved (thus specifying the particular associations for these). By varying these connection strengths, different possibilities for the characteristics and repertoire offered by the modelled agent can be realised. Note that usually weights are assumed non-
negative, except for the inhibiting connections, which are indicated in red colour in the Table 2. For the properties LP: 1, 3, 4, and 5 the function $f$ is taken as the identity function $f(W) = W$ and for all the other states $f$ is a combination function based on the logistic threshold function as in equations (1) (see [29, 30] for more info). In equation (1) $\sigma$ is steepness and $\tau$ is threshold; which are configuration parameters that change the shape of the curve.

$$f(x) = \left( \frac{1}{1 + e^{-\sigma(x-a)}} - \frac{1}{1 + e^{\sigma\tau}} \right) \left( 1 + e^{\sigma\tau} \right) \text{ when } x > 0 \text{ & } f(x) = 0 \text{ when } x < 0$$ \hspace{1cm} (1)

### 4.4. Simulation Results

This section discusses two simulation experiments undertaken to analyse the designed model in different scenarios. In the first scenario it simulates a fight or

**Table 1.** Overview of the connections and their weights. In here the red color $\omega_k$ values for negative weights.

<table>
<thead>
<tr>
<th>from state</th>
<th>to state</th>
<th>weights</th>
<th>LP #</th>
</tr>
</thead>
<tbody>
<tr>
<td>EA(a), EE(e)</td>
<td>WS(s)</td>
<td>$\omega_1$, $\omega_2$</td>
<td>1</td>
</tr>
<tr>
<td>EA(a), EE(e)</td>
<td>WS(b)</td>
<td>$\omega_3$, $\omega_4$</td>
<td>2</td>
</tr>
<tr>
<td>EO(Y)</td>
<td>WS(c)</td>
<td>$\omega_5$</td>
<td>3</td>
</tr>
<tr>
<td>WS(s), WS(e), WS(b)</td>
<td>SS(s</td>
<td>c</td>
<td>b)</td>
</tr>
<tr>
<td>SS(b), PA(a), PO(Y)</td>
<td>SR(b)</td>
<td>$\omega_{11}$, $\omega_{12}$, $\omega_{13}$</td>
<td>6</td>
</tr>
<tr>
<td>SR(s), SR(c), SR(b)</td>
<td>PD(b)</td>
<td>$\omega_{14}$, $\omega_{15}$, $\omega_{16}$</td>
<td>7</td>
</tr>
<tr>
<td>SR(s), PD(b), F(b), PA(e), EAtt(b), EPer(s,b)</td>
<td>PA(a)</td>
<td>$\omega_{17}$, $\omega_{18}$, $\omega_{19}$, $\omega_{20}$</td>
<td>8</td>
</tr>
<tr>
<td>PD(b), SR(b)</td>
<td>F(b)</td>
<td>$\omega_{21}$, $\omega_{22}$</td>
<td>9</td>
</tr>
<tr>
<td>PD(b), SR(c), SR(s), RO(Y), EAtt(b), REA(b,e), SD(b)</td>
<td>EPer(s,b)</td>
<td>$\omega_{25}$, $\omega_{26}$, $\omega_{27}$, $\omega_{28}$</td>
<td>10</td>
</tr>
<tr>
<td>SR(c), SR(s), EAtt(b), REA(b,e)</td>
<td>SD(b)</td>
<td>$\omega_{32}$, $\omega_{33}$, $\omega_{34}$, $\omega_{35}$</td>
<td>11</td>
</tr>
<tr>
<td>SD(b), SR(c), F(b), PEA(b,e)</td>
<td>EAtt(b)</td>
<td>$\omega_{36}$, $\omega_{37}$, $\omega_{38}$, $\omega_{39}$</td>
<td>12</td>
</tr>
<tr>
<td>EPer(s,b), F(b)</td>
<td>PA(e)</td>
<td>$\omega_{40}$, $\omega_{41}$</td>
<td>13</td>
</tr>
<tr>
<td>SD(b), EAtt(b), PA(e), F(b), REA(b,e)</td>
<td>PEA(b,e)</td>
<td>$\omega_{42}$, $\omega_{43}$, $\omega_{44}$, $\omega_{45}$</td>
<td>14</td>
</tr>
<tr>
<td>PA(e), SR(c), PA(a), F(b), RO(Y)</td>
<td>PO(Y)</td>
<td>$\omega_{47}$, $\omega_{48}$, $\omega_{49}$, $\omega_{50}$, $\omega_{51}$</td>
<td>15</td>
</tr>
<tr>
<td>F(b), PO(Y), RAwr(Y)</td>
<td>PAr(Y)</td>
<td>$\omega_{52}$, $\omega_{53}$, $\omega_{54}$</td>
<td>16</td>
</tr>
<tr>
<td>PA(e)</td>
<td>EE(e)</td>
<td>$\omega_{55}$</td>
<td>17</td>
</tr>
<tr>
<td>PA(a), PO(Y)</td>
<td>EA(a)</td>
<td>$\omega_{56}$, $\omega_{57}$</td>
<td>18</td>
</tr>
<tr>
<td>PO(a,b,c,e,s), F(b), EA(a)</td>
<td>RO(Y)</td>
<td>$\omega_{58}$, $\omega_{59}$, $\omega_{60}$</td>
<td>19</td>
</tr>
<tr>
<td>PEA(b,e), EE(e), F(b), RAwr(Y)</td>
<td>REA(b,e)</td>
<td>$\omega_{61}$, $\omega_{62}$, $\omega_{63}$, $\omega_{64}$</td>
<td>20</td>
</tr>
<tr>
<td>REA(b,e), F(b), PAr(Y), RO(Y), EA(a)</td>
<td>RAwr(Y)</td>
<td>$\omega_{65}$, $\omega_{66}$, $\omega_{67}$, $\omega_{68}$, $\omega_{69}$</td>
<td>21</td>
</tr>
<tr>
<td>RAwr(Y), REA(b,e), RO(Y)</td>
<td>EO(Y)</td>
<td>$\omega_{70}$, $\omega_{71}$, $\omega_{72}$</td>
<td>22</td>
</tr>
</tbody>
</table>
flight situation through the bottom-up process [24], and the second scenario simulates the emotion formation in conscious form (with top-down). Selecting suitable weight values for connections in this model was achieved through the same approach explained in [29]. Table 2 lists the connection weight values used for cognitive agent model in the indicated simulation scenarios; threshold (τ) and steepness (σ) values used for those scenarios are listed in Table 3. Furthermore, the step size (Δt) taken is 0.25. The slow value 0.5 for γ was applied for external processes modelled by LP1, LP2, and LP3, and the fast value 0.9 for γ for the internal processes modelled by the other LP’s.

4.4.1. Scenario 1: Fight-or-Flight response

Fast detection and reaction on potential threats are a fundamental adaptation for any being [31] referred as bottom-up in Section 2. The first scenario, shown in Fig. 2 describes a mostly physiological phenomenon called fight or flight response (see [24]) which looks almost automatic as there is no time available for critical cognitive evaluations. Due to the nature of fight-or-flight response (reflexive nature than highly cognitive [15, 16]), and it is an innate process mainly for survival (response time should be relatively low [6] and with a high strength of action execution [15, 16]).

In this fight or flight scenario in the presented model, the context \( c \) is self, and a stimulus \( s \) occurs (which is assumed to have strong emotional associations). As an effect of these inputs in Fig. 2 it is shown that the agent has immediately started to

| Table 2. Connection weight values used for cognitive agent model (Note: all blank cells hold the respective value immediately above that cell). \( \omega \): Weight; \( S \): Simulation. |
|---|---|---|---|---|---|---|---|---|---|
| \( \omega_{1-2} \) | \( \omega_{3-4} \) | \( \omega_{5} \) | \( \omega_{6} \) | \( \omega_{7-8} \) | \( \omega_{9} \) | \( \omega_{10} \) | \( \omega_{11} \) | \( \omega_{12} \) |
| S1 | -0.5 | 0.8 | -0.8 | 1 | 0.7 | 1 | 0.7 | 1 | 0.9 |
| S2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| \( \omega_{13} \) | \( \omega_{14-15} \) | \( \omega_{16} \) | \( \omega_{17-19} \) | \( \omega_{20} \) | \( \omega_{21} \) | \( \omega_{22} \) | \( \omega_{23} \) | \( \omega_{24} \) |
| S1 | -0.9 | 0.9 | -0.7 | 0.7 | 0.9 | 0.7 | 1 | 0.8 | 0.9 |
| S2 | -0.6 | 0.7 | -0.7 | 0.7 | 0.9 | 0.7 | 1 | 0.8 | 0.9 |
| \( \omega_{25} \) | \( \omega_{26-27} \) | \( \omega_{28} \) | \( \omega_{29} \) | \( \omega_{30} \) | \( \omega_{31} \) | \( \omega_{32-33} \) | \( \omega_{34} \) | \( \omega_{35} \) |
| S1 | 0.9 | 0.9 | -0.9 | 0.8 | -0.9 | 0.8 | 0.4 | 0.4 | -0.9 |
| S2 | 0.1 | 0.1 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 |
| \( \omega_{36} \) | \( \omega_{37} \) | \( \omega_{38} \) | \( \omega_{39} \) | \( \omega_{40} \) | \( \omega_{41} \) | \( \omega_{42-45} \) | \( \omega_{46} \) | \( \omega_{47} \) |
| S1 | 0.5 | 0.6 | 0.8 | 0.4 | 0.9 | 0.8 | 0.2 | -0.9 | 0.7 |
| S2 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 |
| \( \omega_{48-50} \) | \( \omega_{51} \) | \( \omega_{52-53} \) | \( \omega_{54} \) | \( \omega_{55} \) | \( \omega_{56} \) | \( \omega_{57} \) | \( \omega_{58} \) | \( \omega_{59-60} \) |
| S1 | 0.7 | -0.9 | 0.4 | -0.8 | 1 | 0.9 | 0.9 | 0.9 | 0.8 |
| S2 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 |
| \( \omega_{61} \) | \( \omega_{62} \) | \( \omega_{63-68} \) | \( \omega_{69} \) | \( \omega_{70-71} \) | \( \omega_{72} \) |
| S1 | 0.4 | 0.4 | 0.4 | 0.5 | 0.4 | 0.9 |
| S2 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 |
develop an emotional perception (mainly based on stimulus $s$, as the preparation for action $a$ has not even got activated yet [15]) around time point 4, with almost non-existing subjective desires [15, 16]. In parallel with the development of the emotional perception, the agent has prepared for action $e$ (emotions) rapidly (around at time point 5) before even the preparation for action $a$ (which is starts around time point 8) as highlighted in [9, 21]. The agent has shown a strong emotional bias having effects from $\text{PA}(e)$ on $\text{PA}(a)$ [15, 16]. This strong emotional bias has led to a strong feeling (which is with the peak value of 0.75) which follows the sensory representation of $b$. The agent has executed the emotional expression of $e$ in a relatively early stage of the timeline (starts around time point 12), and with low emotional attention which got activated relatively late in the timeline [6] (these observations are aligning with the literature on bottom-up process in Section 2). The agent has shown a sufficient strength in prior ownership and subsequently got executed the action $a$ with a very strong peak value: 0.93; this value is the highest peak value observed for the $\text{EA}(a)$ in comparison to the other simulation scenario (see Figure 3). Furthermore, these observations are in line with the explanations of the fight-or-flight response which indicate a tremendous strength in the action [15]. Moreover, it is observed that $\text{EA}(a)$ exists for a considerably longer period of time in comparison to the same in the other scenario (but the overall process time is relatively less [6]). Subsequently the communication of ownership has been followed by a retrospective action ownership with acceptable strength and positions in the timeline (cf. [30]). Note that the agent has not shown any awareness, which is

<table>
<thead>
<tr>
<th>Table 3. Steepness ($\sigma$) and Threshold ($\tau$) values used in configurations of simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simulation One</strong></td>
</tr>
<tr>
<td>$\sigma$</td>
</tr>
<tr>
<td>PD</td>
</tr>
<tr>
<td>SD</td>
</tr>
<tr>
<td>EPer</td>
</tr>
<tr>
<td>EAtt</td>
</tr>
<tr>
<td>PA</td>
</tr>
<tr>
<td>PA</td>
</tr>
<tr>
<td><strong>Simulation Two</strong></td>
</tr>
<tr>
<td>$\sigma$</td>
</tr>
<tr>
<td>PD</td>
</tr>
<tr>
<td>SD</td>
</tr>
<tr>
<td>EPer</td>
</tr>
<tr>
<td>EAtt</td>
</tr>
<tr>
<td>PA</td>
</tr>
<tr>
<td>PA</td>
</tr>
</tbody>
</table>
in line with the evidence from the literature on bottom-up processes, as discussed in Section 2.

4.4.2. Scenario 2: the Top-Down Process

The second scenario presents a simulation on emotion formation through the top-down approach. In this scenario the stimulus may not have a strong emotional association as in the bottom-up process [15]. Therefore mainly through appraisal with a focused intention and subjective desires [14, 15, 28], the agent will experience the emotions and perform the action. An example for this is in [15]: “For example, fear might be elicited from the top-down when someone interprets a curt email from a prospective employer as indicative of disinterest and a low likelihood of being hired.” [15], pp. 254. This simulation is shown in Figure 3; where the context $c$ is the agent itself, and a stimulus $s$ occurs. In Figure 3, part (a) it shows that the agent starts with a performative desire on the given inputs ($c$, $s$), and in part(b) the subjective desires are also becoming prominent (in the timeline PD($b$) are relatively weak and with a short lifespan when comparing with the SD($b$), and that is in line with [15, 16, 23]). Because of the performative desires, the agent triggers preparation of action $a$, which is followed by the sensory representation of the predicted effect $b$ of $a$ (through the internal simulation based on the as-if body
loop) and subsequently by the feeling of \( b \) (with the aid of the activated performative desire for \( b \)) [7, 8]. Primarily because of the predicted feeling, the emotional attention of \( b \) starts to develop (with the influence from subjective desires too [26]) [19, 21, 27, 32]. From this emotional attention of \( b \), the agent starts to develop an emotional perception (primarily on \( b \) in this time) [17], and followed by preparation of action \( e \). Next, these states contribute to generate activation of a prior self-ownership state (cf. [30]). Subsequently the agent develops prior emotional awareness [9, 12, 22, 27, 28] and this leads to the execution of emotional expression (see part (b) in Figure 3). In part (a) of Figure 3, the agent develops prior awareness of action formation and this leads to the execution of action \( a \). By following the emotional expression agent will develop the retrospective emotional awareness and furthermore, after the execution of the action \( a \), the agent will achieve the retrospective ownership, the retrospective awareness, and finally the communication of ownership (cf. [30]). These observations are in line with Section 2.

4.5. Discussion

The computational model introduced in this paper is based on literature from Cognitive and Affective Neuroscience. It incorporates a role for emotional...
awareness states with attention, and perception that act reciprocally and interactively in the dynamics (top-down) of emotion generation, but also covers automatic, unconscious emotion generation processes (bottom-up), and the mutual interaction between these bottom-up and top-down processes [15, 16, 32]. The model was formalised as a dynamical system [4]. Various simulation experiments have been conducted according to different scenarios and the model shows simulation results that are in line with patterns reported in neurological literature. More importantly having two distinct value sets for Steepness (σ) and Threshold (τ) in configurations (for bottom-up and top-down) shows the comparability with the literature where two neural paths also in the human brain for emotion formation [11, 15, 16, 31]. As a summary bottom-up emotions are elicited largely by emotional perceptions with weaker subjective aspects but not necessarily being conscious (reflexive); whereas the top-down is more with conscious and appraisal driven with attention (more cognitive). It is a generic question in this domain how an emotion-laden stimuli processing relates to attention, perception and awareness [17]? The presented agent driven computational cognitive models may further contribute to evaluate, justify and further explore the boundaries with different intuitions to uplift the understanding of the above question. Incorporating a learning mechanism, and processes for emotion regulation will be some future work, together with more validations and comparisons.

References


Chapter 5

Modelling Dynamics of Cognitive Control in Action Formation with Intention, Attention, and Awareness

Dilhan J. Thilakarathne

Agent Systems Research Group, Department of Computer Science, VU University
Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.
e-mail: d.j.thilakarathne@vu.nl

Abstract: Human action formation primarily concerns automatic brain processes that are responsive to a salient stimulus. Nevertheless, the importance of studying the control of these actions to obtain more flexible and self-regulated behaviours under the intervention of top-down related processes has been noted. In this paper a top-down guided action formation based on automatic pathways with the cognitive states intention, attention, and awareness has been modelled. By simulations the validity of the model has been explored. This model will be used in scrutinizing the interplay among conscious and unconscious processes in clinical disorders, as a workbench for cognitive scientists, and in agent-based applications for healthy lifestyle, and complex systems that involve human cognition.

Keywords: intention, attention, awareness, cognitive control, top-down, computational modelling.

---

1 This chapter was published as:
http://doi.org/10.1109/W1-IAT.2014.168
5.1. Introduction

It is a challenge to understand how exactly in the human brain action formation results from the interplay between conscious and unconscious processes. Monsell [1] has stated that as an effect of learning humans have pre-stored the actions per se, and therefore, even in the absence of a particular intention the brain will automatically evoke the relevant action which was habitually associated under the pre-learning process (frequency and recency of a task learned seems directly related to its selection). Nevertheless, even within the same individual the same stimulus (but in different situations), the emotional experiences and performed actions may be different [2]. Furthermore, how humans show flexibility in their behaviour and handling novel situations are difficult to explain merely based on bottom-up driven unconscious processes [3], [4]. Therefore, if the emergence of an action for a stimulus is analogically similar to an orchestra, it is interesting to find whether there is an orchestra behind this [3]. Top-down signals may contribute to the selection of a relevant action by filtering sensory information in a goal-driven guided form. There are evidences that show the importance of top-down (or conscious) factors such as intention, attention, awareness, that are assumed to be playing an important role in this process of cognitive control [4]–[7]. Magic will be an appropriate experience to illustrate the power of cognitive control and its factors. As described by Macknik and colleagues in [8], magicians have the capability to effectively divert (or misdirect) human attention, intention, and awareness in such a way that certain effects seem to be outside the laws of the nature; especially this holds for cognitive illusion techniques, for example, the Vanishing-Ball illusion under inattentional blindness (cf. [8]). They have highlighted that a magician primarily manipulates the spectators’ attention rather than their gaze, through top-down control: effectively they control saliency features of the task by manipulating the audience’s bottom-up and top-down attentional mechanisms [8]. The capacity of the human brain is limited to process all information (cf. [9]), but top-down driven cognitive control may be the innate solution to overcome this limitation.

This paper focuses on a developed computational model that combines both conscious (top-down) and unconscious (bottom-up) processes. The model shows how the interplay of these two in the basis for making the most appropriate response. The model has been inspired from the insights from cognitive, affective, and behavioural sciences presented in Section 2. In Section 3 those evidences have been incorporated in a computational agent model. Section 4 presents an exploration of the model by simulation to get an idea of its validity. Finally, Section 5 is a discussion in which also future directions are pointed out. It may useful to use a model like this as a workbench to conceptually validate some complex phenomenon and/or to further justify their findings and maybe to point at even new directions to conduct experiments. For example, from the clinical perspective,
analysing and validating complex clinical implications related to brain disorders (e.g., autism spectrum disorders (ASD), schizophrenia), and to better understand and treat motor control disorders such as Parkinson’s disease may be some practical applications (cf. [10]). Furthermore, a cognitive model like this can be used to implement agent-based applications that improve human health, especially when cognitive factors are essential.

5.2. Cognitive, Behavioural, and Affective Science Evidence on Top-Down Guided Action Formation

Action formation process is assumed to be automatic, fast and inflexible in general, when the stimulus is “habitual”. Nevertheless, humans have difficulties to continue a reading if the colour ‘red’ printed in ‘blue’ (stroop effect), or to sort cards when the sorting criteria (i.e., shape, colour, number of symbols, etc.) changes periodically (wisconsin card sorting task) [1], [3], [11]. Furthermore, the alien limb syndrome and utilization behaviour are well known neurological disorders that show the importance of using cognitive control on top of the automatic action formation process whenever needed (cf. [12], [13]). For example, patients with the alien limb syndrome, just seeing a cup of coffee might be sufficient to reach and grasp it due to the automatic activation of action plans on this habitual task, irrespective of his/her current location or situation [12] (cf. [10]). Will you do so if that is another person’s coffee?

The prefrontal cortex (PFC) has long been assumed to play an important role in top-down driven cognitive control, as a temporal integrator. The higher order interconnectivity of the PFC with other cortical, and subcortical areas has been interpreted as indicating a process that generates and maintains information when sensory inputs are weak, ambiguous, rapidly changing, novel and/or multiple options exist [3], [14]. Therefore the brain’s circuits for cognitive control seem to consist of loops rather than linear chains (cf. [6]). Furthermore, it has been noted that damages in PFC lead to greater distractibility due to the inability of maintaining a goal relevant focus [14] and this is a reason for disorders such as ASD. In addition, the PFC has been recognized as an important area that is necessary to activate the long term memory stored in the inferior temporal cortex (IT) as discussed in [14].

Intention is also noted as one of the key concepts for conscious action formation [4], [15], and activity in the pre–supplementary motor area (pre-SMA) is assumed to reflect the representation of intention [16] (or intentional actions may relate to a brain network that involving the SMAs (SMA proper and pre-SMA); see [17]). Lau and colleagues [16] have found from fMRI results that when participants attend to their intention, there is enhanced activity in the pre-SMA. Furthermore, by applying transcranial magnetic stimulation (TMS) over the pre-SMA they have found that intention is shifted backward in time while performing a Libet task [17]. Moreover,
they have hypothesized that if a strong activation of the pre-SMA correlates with the awareness of intention, then attention to intention may be one mechanism that contributes to effective conscious control of actions (cf. [6]). The restricted and repetitive behaviour of ASD patients (like "stereotypy" or "perseveration") which is one of the key symptoms of these patients has been interpreted as a reflection of intentional deficits rather than problems in task execution [11].

While it is argued that the unconscious action formation may also involve separate conscious brain processes, strong evidence has been found for an intensive interplay of these two processes, rather than categorically being two independent processes. Especially when inhibiting unwanted unconscious responses through conscious awareness this leads to select the most appropriate response [12]. D’Ostilio & Garraux [10] have highlighted that voluntary control of motor activity involves unconscious initiation, prediction and preparation and then is followed by feeling and awareness of action intention and finally of action awareness. Haynes in [18] has also pointed out that the brain predicts the outcome of a decision even before the decision reaches awareness. Furthermore, Moore & Haggard [4] have investigated whether conscious awareness of an action is based on predictive or inferential processes of action execution (in other terms: awareness of the action and/or the awareness of the effects of an action). They have concluded that awareness of an action is a dynamic combination of both prior awareness (through predictive motor control) and retrospective awareness (through inferential sense-making) relative to the action execution (cf. [4], [13]). The anterior insular cortex (AIC) assumed to play a fundamental role in awareness, accompanied with feelings [19]. The intention to act also has been noted to substantially contribute to awareness of the action [13]. Furthermore, the activity of the prefrontal parietal network (PPN) has been observed in almost all demanding or novel tasks which are usually attributed to cognitive control [20]. An overlap between attention and conscious awareness also has been suggested, due to the PPN which is evoked for the cognitive functions attention, executive functions, and working memory, and also for conscious awareness (this may highlight the interplay among these and possible influences from conscious awareness to attention; though attention might not be necessary and sufficient in order for conscious awareness to emerge; (cf. [20], [21])). Furthermore, functional evidences of top-down mechanisms in attentional selection relating to cognitive control have shown its relation to long term memory and more specifically to the declarative and procedural knowledge in it (cf [5]).

By knowing the limited capacity of the human brain to process all sensory stimuli and/or thoughts present, a cognitive process of information selection (‘attention’) seems to be contributing to focus the processing subjectively and direct towards working memory [9], [22]. Furthermore, for situations where automatic action selection is inadequate it may be necessary to involve executive/voluntary attention.
on action selection to select among competing items, to correct error and to regulate emotions (cf. [23]). The attention process involves bottom-up (mainly driven by salient features of external stimuli) and top-down factors (by prior knowledge, wilful plans, intentions, behavioural relevance) [9], [21], [22]. In [9] it has been pointed out that the posterior parietal cortex (PPC) and PFC, could be segregated for distinct roles in bottom-up and top-down attentional systems, and the close interaction of these regions with each other is highlighted to explain the constant influence of these two processes to orient the attention as necessary for more sophisticated cognitive control process (cf. [11], [20], [22]). In addition to the intention, PPC is also assumed to be involved in spatial attention, spatial awareness, polysensory integration, and decision making [24]. While the prediction and selection of intentional actions (in unconscious manner) have been assigned to the activity of the PPC, the intention has been related to the SMAs [17] thereby showing the relation of these two factors. Furthermore, once the focus of attention is in a location, if another stimulus appears in the same location, suppression has been reported in the PPC which may be interpreted in the sense that the current attention may have suppressed the emergence of new attention (cf [9]). Knudsen in [25] has given four processes, working memory, top-down sensitivity control, competitive selection, and automatic bottom-up filtering for salient stimuli, as functions for attention (top-down attention has been attributed to the first three as a recurrent loop, and specifically for the bottom-up attention see [7]). Tallon-Baudry has highlighted three hypotheses for possible relations in attention and consciousness by relating to different experimental findings: gateway, reverse dependence, and cumulative influence (cf. [21]). It has been found that the primary motor cortex (M1) region is also modulated by top-down effects of attention which gives some intuition about the relation from attention to motor preparation [17].

Inhibition mechanisms may also be as important as the excitation mechanisms in cognitive control (though some different viewpoints also provided (see [26])). By using Gamma-aminobutyric acid (GABA), neurons are performing inhibition at synaptic, circuit, and systems levels (cf. [26]). Furthermore, various inhibition types in neuroscience and psychology have been discussed in [26]; inhibition also activates in automatic (e.g., lateral inhibition: if a particular representation accumulates more evidences, that will suppress its fellow representations) and voluntary manner (e.g., suppression of an irrelevant response, stimulus, or memory; in intentionally). Furthermore, [25] highlights the inhibition circuitry associated with top-down attention: inhibiting the neurons that are not tuned for the current selection, and exciting the activation of a selected path to tune, makes that a balanced influence by the excitatory and inhibitory circuitry in order to accomplish a sharp feature tuning.

Emotions also play a vital role in action formation processes; their contribution has been pointed out mainly on two paths: perceptual and cognitive control [27].
The perceptual path is integrated with bottom-up process, whereas cognitive control is with the top-down process. More information, specifically about bottom-up and top-down processes in emotion generation, and their mutual interaction together with emotional awareness can be found in [28].

5.3. Description of the Computational Model

With the literature highlighted in the above section, a computational cognitive model will be presented here focusing on the dynamics of cognitive control in action formation with intention, attention, and awareness. Furthermore, the associated circuits for cognitive control in the human brain consist of loops rather than linear

![Diagram](image-url)

**Fig. 1**: Overview of the computational cognitive agent model. Here; Y: a, b, c, e, s.

The arrow represent a direct activation to state B from state A, arrow represent a direct suppression to state B from state A, arrow represent a suppression to all the complements of ‘ith’ state on B, from state A (where ‘i’ presents an instance of a particular state), and represent a direct activation to state B, from state A while suppressing all the complements of ‘ith’ state on A, from
chains. Therefore, it will be reasonable to use modelling techniques for this cyclic connectivity. The model proposed here adopts and integrates fragments from previously developed cognitive models: for ownership of actions [29], awareness of actions [30], intentional inhibition of actions [31], and emotional awareness of actions [28]. But the newly introduced model focuses in particular on cognitive control in human action formation. An overview of the proposed model is shown in Fig. 1 and its abbreviation details can be found in Table 1. Primarily this model covers two causal action formation processes: bottom-up and top-down. The bottom-up process covers the action formation from an unconscious perspective and the top-down process from a conscious perspective.

5.3.1. Action Formation Unconsciously

The model uses three world states (WS) as inputs: for stimulus s: WS(s), context c: WS(c), and effect b: WS(b). The stimulus s represents any internal (bodily: e.g. self-generated facial expression) or external change that may lead to an action execution. Context c represents additional information perceived to improve the process of action selection (c can be differentiated as ‘self’ and ‘other’ as explained in [29], [30], and effects of mirror neurons have been explained when c is ‘other’). The effect b represents the effects of the execution of an action a. The input world states WS(s), WS(c), and WS(b) lead to sensor states SS(s), SS(c), and SS(b), and

<table>
<thead>
<tr>
<th>WS(W)</th>
<th>world state W (W can be either: context c, stimulus s, or effect b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS(W)</td>
<td>sensor state for W</td>
</tr>
<tr>
<td>SR(W)</td>
<td>sensory representation of W</td>
</tr>
<tr>
<td>PD(b)</td>
<td>performative desires for b</td>
</tr>
<tr>
<td>SD(b)</td>
<td>subjective desires for b</td>
</tr>
<tr>
<td>PA(e)</td>
<td>preparation for emotional response e</td>
</tr>
<tr>
<td>PA(a)</td>
<td>preparation for action a</td>
</tr>
<tr>
<td>F(b)</td>
<td>feeling for action a and its effects and emotions</td>
</tr>
<tr>
<td>Per(s,c,b)</td>
<td>perception state for s and c on b</td>
</tr>
<tr>
<td>Att(b)</td>
<td>attention state for b</td>
</tr>
<tr>
<td>CInt(b)</td>
<td>conscious intention state for b</td>
</tr>
<tr>
<td>EE(e)</td>
<td>expression of emotion e</td>
</tr>
<tr>
<td>EA(a)</td>
<td>execution of action a</td>
</tr>
<tr>
<td>PO(Y)</td>
<td>prior ownership state for action a with b, c, e, and</td>
</tr>
<tr>
<td>RO(Y)</td>
<td>retrospective ownership state for a with b, c, e,</td>
</tr>
<tr>
<td>PEA(b,e)</td>
<td>prior emotional awareness state for b with e</td>
</tr>
<tr>
<td>REA(b,e)</td>
<td>retrospective emotional awareness state for b</td>
</tr>
<tr>
<td>PAwr(Y)</td>
<td>prior-awareness state for action a with b, c, e, and</td>
</tr>
<tr>
<td>RAwr(Y)</td>
<td>retrospective-awareness state for action a with b,</td>
</tr>
<tr>
<td>EO(Y)</td>
<td>communication of ownership of a with b, c, e,</td>
</tr>
</tbody>
</table>

Table 1: Nomenclature for Fig. 1
subsequently to sensory representation states SR(s), SR(c), and SR(b), respectively. The unconscious causality of action formation has been modelled as explained in [32]: by combining Damasio’s as-if body loop (see [33]) and the James’s body loop (see [34]) hypotheses. James’s body loop consists of: stimulus → sensory representation → preparation for bodily changes → expressed emotion → felt emotion. This has been mapped in this model to causal relationships between: WS(s), SS(s), SR(s), PA(a), EA(a), WS(b), SS(b), SR(b), and F(b) (see Fig. 1). Damasio extended the body loop concept and argued that the cognitive process of action selection is due to an effect of an internal simulation process prior to the execution of an action. The brain will evaluate the effect of each relevant action option (i.e. PA(a_i)) by comparing feelings associated to their individual valuated effects (without actually executing them through the body loop). The simulated option that has the strongest (positive) valuated feeling performs as a GO signal through the body loop and else are NO-GO options. The as-if body loop consists of: sensory representation → preparation for bodily changes → felt emotion [33]. In this model this is represented by the loop through: PA(a), SR(b), and F(b). Furthermore, in Fig. 1, state labels are attached with subscript letters: k and i; which indicates for example the kth instance for a stimulus s (e.g. WS(s_k)) and the ith option for an action a (e.g. PA(a_i)). In this model each PA(a_i) state uses its valuated feeling of that option (see Fig. 1) and proportional to the accumulated strength of the state, it suppresses its complementary options on a_i for all PA(a_j) with j≠i (as shown in dotted looped red arrow in Fig. 1). Therefore, naturally the strongest internally satisfied (exceeding a threshold value) option will be getting selected. Furthermore, this bottom-up process involves performative desires PD(b), as explained in [31] to facilitate short-term desire effects on action execution.

In this model to express emotions (complex reactions to certain stimuli by the body [33]), there is an extended route available: Per(s,c,b) → PA(e) → EE(e). This is mainly obtained from the bottom-up emotion generation literature: emotions can be shaped by the perception through amplification mechanisms that may not necessarily overlap with attentional processes [35]. Therefore, in this model PA(a) gets effects from PA(e) (more information on this can be found in [28]). Similar to the attention-related modulation mechanism in Section 2 (see [9], [25], [26]), this also includes the same process, that is while Per(s_k,c_k,b_i) is exciting the preparation activation of option e_i: PA(e_i); it suppresses all the complementary preparations of e_i: all PA(e_j) with j≠i. The feeling state (subjective experience of emotions [33]) also plays a vital role in different ways in this model. The activation levels of the states PA(a), PA(e), PO(Y), Att(b), PEA(b,e), PAwr(Y), REA(b,e), and RAwr(Y) get effects from F(b). The main reason for these effects lies in the as-if body loop performing the valuation based on the feeling. Therefore the higher order states (including conscious-related) also may trigger through this (cf. [33]). Damasio has attributed emotions more into
the unconscious processes and feelings as a basis for conscious processes. In this model the feeling state has been modelled as an important element for both conscious and unconscious processes but on a very strong stimulus which has salient features, still it is possible to directly execute the action through the above mentioned emotion driven route (Per(s,c,b), PA(e), and EE(e)) rather than internally simulating the option through the as-if body loop (in [28], this ability of executing a quick and very strong action has been modelled and simulated mainly from a suddenly developed emotional state through a strong perception triggered from a stimulus under the bottom-up approach for a situation called fight-or-flight).

Together with action preparation processes, ownership of an action is developed: in how far does a person attribute an action to him or herself or to another person; this also gets triggered as explained in [29]. The prior ownership state PO(Y) is mainly affected by the predicted effects of prepared action (also from: F(b), SR(c), and PA(e)) and functions to control the actual execution, as specifically explained in [29]. Based on the predicted feeling (for the effects), an automatic suppression of the sensed actual effect will take place but as a combination of these two effects (suppression and activation) RO(Y) will be developed and based on that also acknowledged authorship of the executed action through EO(Y) (cf. [4]). Therefore, there is a suppressive effect on SR(b) from PO(Y) and this preliminarily contributes to separating prior effects from retrospective effects.

5.3.2. Top-Down Controlling Effect on Action Formation

Ownership states were attributed to the unconscious action formation. From the other side intention is a conception formed for a goal-directed preparation or movement or determination, attention concerns selecting or focusing one or few from a cognitive content, and awareness concerns his or her subjective experience of a cognitive content. All these play their role in conscious processes of action formation. As explained in Section 2, attention can be bottom-up attention, top-down attention or a combination of the both [9], [21], [22]. In this model attention is affected by predicted feeling F(b), aligned with bottom-up attention evidences. Top-down attention modulates the preparation of action option ai: PA(ai) (i.e., increasing the activation of option ai while suppressing all its complementary options PA(aj) for all j≠i). The bottom-up attention passes properties (or salient features) of sensory stimuli for higher order cognitive processes, and top-down attention effort fully influences the action formation process; this causes some bias effects in the processing. Due to this cyclic dependency of bottom-up and top-down attention processes in this model, it is in line with the literature evidence of a combined effect of these two (see Fig. 1) [11], [20], [22]. Furthermore, computationally it has been assumed that the suppression by one option bj of its all complements bi with j≠i has been always proportional to the strength of the selected option. For example, if there
are two options $b_1$, $b_2$ and strength of $b_1$ > strength of $b_2$, then the suppression from $b_1$ on $b_2$ should be relatively higher compared to the suppression from $b_2$ on $b_1$ (i.e., $\text{Att}(b_1)$ suppresses $\text{PA}(b_2)$ much stronger) [25]. This demonstrates the focusing nature of the attention process as highlighted in [9], [22]. Dynamically a cognitive overloaded situation may be transferred to a focused, biased and subjective situation (which will be useful to explain the effects in stage magic highlighted in [8]). Additionally it has been found that a direct association between attention and awareness of emotions and having higher emotional awareness will always contribute to accurately detecting and discriminating emotional signals [36] (for more information see [28]). Therefore, there is a bilateral connection between $\text{Att}(b_i)$ and $\text{PEA}(b_i,e_i)$ in this model (emotional awareness effects also have been separated into prior and retrospective effects, by analogically considering the evidences presented in [4]). Emotional perception was also found to be modulated by attention from the evidences in [37], and therefore in this model this has shown the effects of $\text{Att}(b_i)$ on $\text{Per}(s_k,c_k,b_i)$.

Subjective desires (or constitutive desires) $\text{SD}(b)$ are also important in top-down processes to regulate certain behaviours (usefulness of this feature has been modelled in [31] under intentional inhibition). Having this state $\text{SD}(b)$ may contribute to assign alternative meanings (or interpretations) on proximate desires based action preparations and direct the attention to a different state through the intention [5]. Therefore, this model contains a cyclic route from attention, to subjective desires to intention so that less salient features obtained through bottom-up attention can be subjectively controlled in a more conscious manner. Furthermore, through the intention state (i.e. $\text{CInt}(b_i)$), it is even possible to consciously initiate the action formation process due to having direct effects from both $\text{SR}(s_i)$ and $\text{SR}(c_i)$ [13] (this may be useful on the restricted and repetitive behaviour of ASD patients as in [11], who are unable to stop their abnormal behaviour, though they can clearly observe this misbehaviour, due to a deficit in intention). Intention state is also modelled with the same modulation process (on action preparation state) as explained for attention on action preparation. Intention state: $\text{CInt}(b_i)$; further suppresses complementary options of $b_i$, so that this will further strengthen the top-down driven controlling.

Awareness states are among the most challenging and debatable concepts in cognitive modelling. Two forms of awareness states are available in this model to improve the clarity of the model on literature evidences [4], [13], [27]: emotional awareness $\text{PEA}(b,e)$, and $\text{REA}(b,e)$, and more abstract awareness $\text{PAwr}(Y)$, and $\text{RAwr}(Y)$. The prior and retrospective awareness are modeled as suggested in [4]: based on predictive (i.e. as-if body loop) and inferential (i.e. body loop) processes of action execution. Furthermore, as mentioned in Section 2, awareness of action is a dynamic combination of both prior awareness and retrospective awareness relative to the action execution [4], [13]. This model includes effects of $\text{PAwr}(Y)$ on $\text{RAwr}(Y)$
and shows that execution of action $EA(a)$ may directly have effects from prior awareness. The effect of intention on awareness was highlighted in [13] and is also included in this model. Therefore, prior awareness can influence the action execution further. Additionally, due to the overlap between attention and awareness highlighted in [20], [21], this directional dependency is also included in this model (as explained under the attention process earlier). Finally, according to the information provided in [10], [18], the emergence of awareness states should follow the action prediction processes. Therefore, in the timeline of action formation process it is important to show this phenomenon and having the awareness states in a higher level of this model this should be achievable.

5.3.3. Dynamics of the Model

Every connection between state properties in Fig. 1, has a weight value ($\omega_{ji}$: weight of state $j$ to $i$) as a parameter. A weight always has a value between -1 and +1. For different situations with specific context $c_i$, stimulus $s_i$, action $a_i$ and/or effect state $b_i$ involved, it is necessary to isolate the proper weight value vector to simulate that particular situation (or in other words, by varying these connection strengths, different possibilities for the characteristics and repertoire offered by the modelled agent can be realised). In general, weight values are assumed to be non-negative except when it is a suppressive (or inhibiting) link. If it is assumed that $k=1$ and $i=1$ in this model, then there will be 83 weight values (this will increase rapidly when $i = 1 \ldots n$ and further when changing $k = 1 \ldots m$). To model the dynamics following the connections between the states as temporal–causal relations, a dynamical systems perspective is used as explained in [32]. For this purpose each state includes an additional parameter: speed factor $\gamma_i$ indicating the speed by which an activation level is updated upon received input from other states to the state ‘$i$’. Mainly two speed factor values used: slow and fast; in which slow speed factor value was attached to the external states: $WS(s)$, $WS(c)$, $WS(b)$, $SS(s)$, $SS(c)$, $SS(b)$, $EE(e)$, $EA(a)$, and $EO(Y)$, and the rest (internal states) are with fast speed factor value. Activation of a state is depending on multiple other states that are directly attached to it and therefore incoming activation levels from other states are combined to some aggregated input and performed the activation according to a differential equation as in (1) (where $f$ is a combination function on aggregated inputs, and $y_i$ is the activation level of state $i$).

$$\frac{dy_i}{dt} = \gamma_i \left[ f \left( \sum_{j \in \partial(i)} \omega_{ji} y_j \right) - y_i \right] \quad (1)$$

In this model for combination function a continuous logistic threshold function is used as in equation (2) and when the aggregated input is negative (3) is used (where $\sigma$: steepness, and $\tau$: threshold).
Further to achieve the temporal behaviour of each state as a dynamical system, a difference equation is used in the form of equation (4) (where Δt is the time step size). More details about all these equations can be found in [32].

\[
f(X) = \begin{cases} 
\frac{1}{1 + e^{-\sigma(X-\tau)}} - \frac{1}{1 + e^{\sigma \tau}} & \text{when } X > 0 \\
0 & \text{when } X \leq 0
\end{cases}
\]  

\[
f(X) = 0 \text{ when } X \leq 0
\]

There are some inhibition links in this model mainly for the computational purpose to stop the stimulus \( s \) and/or context \( c \), and also to slow down some prior processes after executing the action. Therefore WS(s) gets inhibition effects from the both EA(a) and EE(e), WS(c) gets inhibition effects from EO(Y), SR(b) gets inhibition effects from RO(Y), PO(Y) gets inhibition effects from RO(Y), PEA(b,e) gets inhibition effects from REA(b,e), and PAwr(Y) gets inhibition effects from RAwr(Y).

5.4. Analysis of the Model Based on a Simulation Experiment

The proposed model in Section 3 with the literature highlighted in Section 2 (including the contents of previous works [28]–[31]) was validated considering a scenario and simulating its behaviour. In this scenario only a single instance of a stimulus \( s \) is used together with a single context \( c \) (i.e. \( k=1 \)), and these inputs will trigger two action options (i.e. \( i=\{1,2\} \)). More importantly the focus of this simulation is to highlight the top-down driven cognitive control with attention, intention, and awareness, and therefore it is made sure that when disabling these states (and the related other conscious states) both options have the same configurations (weight values, and the remaining parameter values). Moreover, if the individual option is executed (i.e., in the first simulation \( k=1 \) and \( i=1 \); and in the second \( k=1 \) and \( i=2 \)) the identical executions are obtained for both options individually. Therefore, in unconscious action formation stage there is no bias or salient features with any option and through the as-if body loop both options will be selected as GO signal. This scenario simulates how the top-down processes contribute to selection of a suitable option and the temporal nature of this action formation, by providing intention, attention and subjective desires on option \( i=1 \).

The above mentioned model was developed in Java as a framework together with relevant other technologies to handle inputs (in XML format) and to visualise the behaviour in graphs. Inputs to the model were:

- step size (Δt) 0.25
• Slow speed factor ($\gamma$) 0.6
• Fast speed factor ($\gamma$) 0.9
• Total time slots 200 (number of execution epochs on given step size)

Moreover, steepness ($\sigma$), threshold ($\tau$), and weight values for each state were specified separately. All input values, difference equations for each state, and enlarged output graphs; are included in a separate appendix2.

Fig. 2 and Fig. 3 includes output patterns for the above mentioned scenario (here two graphs have been provided merely for the clarity of images but information of the both graphs obtained from a single execution where two options existed and only the first option emerged as the action execution due to the cognitive controlling). Fig. 2 shows that the action formation processes starts with a weak proximate desire (PD($b_1$)) (max ~0.2) which gives the intuition that short term desires are not prominent for this scenario, nevertheless a strong action preparation PA($a_1$) is observed (max ~0.8). When considering the same in Fig. 3 it shows that both PD($b_2$) (max ~0.24) and PA($a_2$) (max ~0.28) with low values, but specifically PA($a_2$) having such a small value it highlights the top-down effects on $a_2$. Together with the action preparation PA($a_1$), the sensory representation SR($b_1$) of $b_1$ is activated as explained by the as-if body loop and more importantly there is a dip in this curve (close to the time action execution occurs) and then it will increase after action is executed. This observation in the simulation is aligning with the explanation in [4], where it is described that the predicted effect and the sensed

---

actual effect are added to each other in some integration process (cf [29]). Furthermore, this observation provides clear evidence for the prior and retrospective effects as highlighted in Section 2 and contributes for a clear interplay between predictive processes and actual sensed processes. Furthermore, in parallel to this, the feeling $F(b_1)$ also triggered and shows positive support compared with $F(b_2)$ in Fig. 3.

Together with the feeling, the predictive process in Fig. 2 has shown the activation of attention $Att(b_1)$ and getting it strengthened due to the mentioned top-down driven attention process. The state $Att(b_2)$ shows very low activation (max 0.08) in Fig. 3 and therefore it illustrates the bias nature of cognitive control ($Att(b_1)$ on $Att(b_2)$) as highlighted in [9], [11], [20], [22]. Together with $Att(b_1)$, the subjective desires $SD(b_1)$ of $b_1$ (max $\sim$0.52) and the conscious intention $Clnt(b_1)$ of $b_1$ (max $\sim$0.7) is also activated in Fig. 2 thus highlighting the cycle explained in Section 2. Furthermore, the perception $per(s_t,c_t,b_1)$ is relatively very small compared with the attention in Fig. 2 as the input stimuli do not contain strong salient features to develop strong perception at a very early stage of the timeline (cf. [28]). Subsequently, preparation for emotions $PA(e_1)$ also developed with an adequate strength (max $\sim$0.45). When considering the behaviour of all these states in Fig. 3 it is very clear that none of them is able to sufficiently develop due to the highlighted bias nature.

From the influence of the above mentioned cognitive states, prior ownership $PO(Y_1)$, prior emotional awareness $PEA(b_1,e_1)$, and prior awareness $PAwr(Y_1)$ has been developed in the mentioned order (see Fig. 2) and more importantly as suggested in [4], [10], [13], [18]. Therefore, it is very clear from this simulation data

---

**Fig. 3:** Action formation for option 2.
that the awareness has been fully developed after the predictive processes but just
closer to action execution. With the mentioned awareness states expression of
emotions $EE(e_i)$ and execution of action $EA(a_i)$ has fired respectively in Fig. 2.
Nevertheless, in Fig. 3 none of these have occurred and therefore it is not possible to
see any retrospective effects in that particular to option 2 ($i=2$). In Fig. 2 after
executing the action the retrospective ownership $RO(Y_i)$, the retrospective emotional
awareness $REA(b_i,e_i)$, the retrospective awareness $RAwr(Y_i)$, and finally the
communication of ownership $EO(Y_i)$ occur. The retrospective awareness related
behaviours aligned with [4], [13] and further these outcomes are consistent with the
results obtained in previous models in [28]–[31] as well.

5.5. Discussion

The patterns observed in the previous section give confidence that the proposed
model in cognitive control in action formation agrees with the highlighted literature.
This model shows the interplay between the bottom-up and the top-down processes
and how they interact with each other to cognitively drive a given situation to a
probable solution with attention, intention, and awareness. Including the aspects of
performative and subjective desires, perception, emotion, feeling, ownership,
attention, intention, and awareness, this model has adequate levels of detail (or
complexity) to be used as a coherent system for various experimental needs to
analyse behaviours (may be beneficial as a work bench for some hypotheses in
cognitive, affective, and behaviour sciences). Furthermore, it is easily possible to
aggregate some of these states and to transform this model into a compact version
based on the requirements to reduce the complexity of the model. In cognitive multi-
agent applications this model may be useful to be used as an agent’s working
knowledge and to simulate behaviours both individually and as a community. This
model further needs to be validated with many scenarios as future work. For that
from clinical perspective cognitive disorders like ASD, ADHD, schizophrenia, and
motor control disorders such as the Parkinson’s disease will be good options and for
AI this may be useful to develop healthy lifestyle apps, and to scrutinize situation
awareness in domains like air traffic management, energy management of a small
scale grid when human agents are involved; for such domains it will be a good point
of departure.

Acknowledgment

I wish to thank Prof. Jan Treur at VU University Amsterdam, for his great
support and supervision in all the phases of this work.
References

[34] W. James, “What is an Emotion?,” Mind, vol. 9, no. 34, pp. 188–205, 1884.
Part III:

Modelling Cognitive Metaphor in Joint Decision Making

Making decisions together with others is an essential part of human life, especially in social and professional context. The way we understand many phenomena in our daily lives is metaphorical: one certain mental domain is understood in terms of another mental phenomenon. It is interesting to explore what can be learnt from cognitive metaphors for joint decision making in particular. Therefore, the influence of cognitive metaphors on joint decision making has been examined by combining the concept of a joint decision making process and cognitive metaphor together in a computational social agent model. In close relation to this, more in general, understanding new things based on what already has been understood is also essential, which is referred as analogy making. Analogy making is a fundamental process of human cognition. Therefore, this research exploits the Buddhist theory of Five Aggregates to model analogy making. This shows the possibility of adopting different types of literature in cognitive modelling.
Chapter 6
Scrutinizing the Use of Emergent Intelligence for Analogy Making through Five Aggregates

Dilhan J. Thilakarathne
Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.
e-mail: d.j.thilakarathne@vu.nl

Abstract: Analogy making is a fundamental process of human cognition that helps to understand new things based on what already has been understood. Even though this is an innate ability of human beings, embedding this feature into a computational model is a research challenge. This research exploits the Buddhist theory of Five Aggregates to model analogy making. The Five Aggregates comprise of what is known as Form, Sensation, Perception, Mental formation and Consciousness. Form is the input to the mind while sensation is the basic acceptance of that input into the mind. Perception is about the identification of the input via previous knowledge. Mental formation is about reasoning, which emerges by meeting the conditions that stem from the already gathered knowledge. Here, consciousness is nothing but the ready-state of the mind. The approach was evaluated for a geometric domain.

Keywords: Analogy, Emergent Intelligence, Five Aggregates, Conditioned Phenomena
6.1. Introduction

Through the evolutionary process of Mother Nature, one of the fascinating features that was developed in human beings is intelligence (Rotha et al., 2005). Western philosophy attempts to understand intelligence going back to the ancient Greeks; throughout centuries many philosophers have tried to define the tacit term intelligence (Cianciolo et al., 2008). Various definitions, thoughts, and insights have been suggested for intelligence with mere objective senses (Legg et al., 2007); nevertheless, there is no, well specific definition for intelligence.

Living can often be considered as reasoning about new problems based on familiar, but distinct concepts that have been already understood (French, 2002). By doing so, humans develop new knowledge and further extended this process as a learning process. Terms such as ‘using metaphors’ and ‘analogy making’ refer to such processes, which are fundamental cognitive processes (Hoffman, 1995; Holyoak et al., 2001; Thagard et al., 1990). Henry David Thoreau, the philosopher wrote in his journal in 1851, “All perceptions of truth is the detection of an analogy; we reason from our hands to our head” (Thoreau, 1962). Therefore, though it is not precisely sure what intelligence is, it undoubtedly and almost inevitably reveals that an important element of human intelligence is shown by the skills of analogy making of that person.

6.2. Motivation

American poet John Godfrey Saxe (1816-1887) wrote a poem “The Blind Men and the Elephant”; that gives the rationale behind analogy making and its associated complexity as a perceptual process. Analogy making is the key mechanism for us to perceive nature, though it is a fabrication most of the time. It is quite interesting when children refer to the bright sun, as a big red or yellow ball. In the same manner, an undergraduate physics student who studies light realizes it by comparing and contrasting the behaviour of water waves and marbles that give the idea of wave particle duality. Furthermore, even if we have not seen an atom, universe, deoxyribonucleic acid, cancer cell, beyond three dimensions, et cetera, in science those are well known conceptual innovations allowing us to attempt to solve a numerous number of phenomena successfully (Nersesian, 2008). Viewed from a theoretical perspective, the theory of relativity, quantum theory, unified field theory, big-bang theory, and string theory are nothing but merely analogies formed through a cognitive process. Viewed from an engineering perspective we have developed an aero-plane, submarine, house analogically, to a bird, fish, and cave, respectively. In general, we are an analogy making species.

It seems that we cannot explain anything in this world as it is. If we want to explain anything, we simply relate it to something which is in our ontology.
Knowledge or the long-term memory is what is known as our ontology in this regard, which has ideas with labels that we can retrieve without questioning. For an example, if I ask the question ‘do you know who is a mother’, the answer is obvious; nevertheless, the answer for ‘who is a mother’ is purely an outcome of an analogy made. Analogies characterize a particular thought (that can be a phenomenon, situation, concept, for example) from related thoughts. Analogy making is an innate ability of humans and is the core of human cognition (Holyoak et al., 2001).

From the early developments of the Artificial Intelligence (AI) discipline, analogy solving is considered a micro-world problem (Russell et al., 2009); the first analogy program was presented in 1964 (Evans, 1964). There are significant developments and improvements from that research which are explained in Section 6.3. In this research, analogy making has been approached through emergent intelligence as a conditioned phenomenon. Intelligence can be considered as an emergent property (Rzevski et al., 2007); according to the Buddhist philosophy; everything is a conditioned phenomenon (Jayatilleke, 1978). The cognitive model considered for this is given in Section 6.5.

An agent model was developed that was enriched with an abstract model created out of five aggregates (Piyadassi, 2008) used to express a ‘being’ It has been modelled as an analogy solver in Section 6.6. Emergent behaviour has been demonstrated through the communication and negotiation abilities of agents by conditioning via interaction. This approach was evaluated for a geometric domain in Section 6.7.

6.3. Evolution of Analogy Making Models

Modelling an analogy solver is not trivial. The first analogy program in history was presented in 1964 by Thomas G. Evans (Evans, 1964), and various theories and models have been developed since then. This section presents the development of analogy modelling approaches, with their strengths and limitations.

6.3.1. Structure Mapping Theory Related

Structure Mapping Theory (SMT) by Dedre Gentner in 1983, is considered as one of the strongly recognised works on analogy making. It has had a strong influence on other analogy making models, namely, Structure Mapping Engine (SME) (Falkenhainer et al., 1986), Greedy Structure Mapping Engine (GSME) (Forbus et al., 1990), ISME (Forbus et al., 1994), Many Are Called but Few Are Chosen (MAC/FAC) (Gentner et al., 1991; Forbus et al., 1995).

Gentner pointed out that analogy can be seen as a mapping of knowledge between two domains, merely through same system of relations that hold between source and target domains, and then feature a mapping between these systems
(Gentner, 1983). As Gentner believed, “An analogy is a comparison in which relational predicates, but few or no object attributes, can be mapped from base to target” (Gentner, 1983). Furthermore, he pointed out that higher order relations constrain the lower-order ones due to the central idea of SMT: “principle of systematicity”. In his theory, the priority is always for the structure but not for the content. SMT contributes to many computational models of analogy, metaphor, case-based reasoning and machine translation (Veale, 1998).

SMT is useful as it ignores surface features and finds matches between potentially very different things if they have the same representational structure. Nevertheless, this approach brings out a NP-Hard problem due to the principle of systematicity, as we want to find higher order relations which are subjective (Veale, 1998). Therefore, in practice it is clear that we may get more than one solution for a given analogy problem. However certain applications of SMT have pinpointed the usefulness of it in various models.

6.3.2. Analogical Constraint Mapping Engine

The approach called Analogical Constraint Mapping Engine (ACME) was created by Holyoak and Thagard (Holyoak et al., 1989). It was the first connectionist satisfaction network of an analogy making model, and it is somewhat homogeneous with SME due to use of structural (configuration-related) constraints to interpret a given analogy, distinguishing semantic (meaning-related) and pragmatic (goal-related) influences in the model. SME was restricted to one-to-one mappings but ACME uses many-to-one mappings that fully allow pragmatic and semantic constraints, in addition to the structural constraints (Krawczyka et al., 2005). ACME’s architecture gave an emergent result from constrained parallel activation of states in a neural network-like structure (French, 2002).

Connectionist systems can only represent knowledge implicitly due to the use of weights in connected links, and also due to the power of adaptation, novel solutions may emerge too. In ACME, structural constraints are supported to strengthen the relational correspondences between elements of source and target domains whilst semantic similarity is supported by giving similar meanings within the matches (Holyoak et al., 1989). Pragmatic influence ensures that elements that are involved in analogy making believed to be more important in order to achieve a high quality solution. However, ACME also does not guarantee to provide the best solution by including all the important matches when the number of predicates is more than twenty (Holyoak et al., 1989).

6.3.3. Analogue Retrieval by Constraint Satisfaction

Analogue Retrieval by Constraint Satisfaction (ARCS), by Paul Thagard and co-workers was an extended version of ACME model (Thagard et al., 1990). The main
problem that was noted in analogy retrieval was how to get only what is expected from memory without retrieving anything else than what is to be used (Thagard et al., 1990). ARCS also use the structural, semantic, and pragmatic constraints as similar to ACME, but their relative importance has changed. ACMS considers that semantic similarity constraints have a higher importance in retrieval; therefore, ARCS also has two stages similar to the MAC/FAC, one merely for retrieval (semantically similar domain retrieval) and next for ACME based mapping.

Furthermore, the authors have stated three important challenges in retrieval for any successful computational analogy retriever; namely “It must efficiently find relevant analogues”; “The system must be able to screen out analogues that are less relevant”, and “If the system is intended to be a psychological model, the analogues it retrieves and the processes by which it does so should correspond to human performance as exhibited in controlled psychological experiments” (Thagard et al., 1990). ARCS takes a considerable amount of time to make analogies due to its multiple stage of searching on huge data sets. Furthermore, ARCS has a limitation when more than one solution is required for a new problem (Boden, 1992).

6.3.4. Incremental Analogy Machine

Incremental Analogy Machine (IAM) constructs analogies in an incremental fashion by combining mapping and the retrieval models SME and ACME (Keane et al., 1988). According to Keane and co-workers, the IAM algorithm is as follows: 1. Select Seed Group, 2. Find Seed Match, 3. Find Isomorphic Matches for Group, 4. Find Transfers for Group, 5. Evaluate Group Mapping, and 6. Find Other Group Mappings (Keane et al., 1994).

IAM’s incremental process maps a portion of its target domain to the source domain which incrementally leads to a single solution, and evaluates its strength (French, 2002). If it is not optimal, it will loop again with another portion of its target domain. This is a purely sequential approach; parallel alternative interpretations are not supported (French, 2002). Furthermore, IAM is a mapping model, somewhat cognitively plausible. Due to this nature, it does not have performance issues that are noted in SME and ACME, as through the Seed Group idea with identification of root predicates, it can handle a large amount of data successfully and allow retrieval of more than one analogy too. Nevertheless, the completely serial nature of IAM processing has produced doubts about its ability to scale up (French, 2002).

6.3.5. Learning and Inference with Schemas and Analogies

Learning and Inference with Schemas and Analogies (LISA) is a neural network model (connectionist model) that enables the relationship between analogy and schema induction (a schema is a knowledge structure) (Hummel et al., 1996) by
combining the advantages of the symbolic and sub-symbolic approaches to cognitive modelling. LISA is based on synchronic activation for temporal dynamic binding. Memory access based retrieval and structural mapping were the main concerns of this analogy making model. The model contains two interacting systems namely Working Memory (WM) and Long Term Memory (LTM). LISA represents both objects and relational roles in a distributed fashion as propositional representations in LTM (Hummel et al., 1997).

The process is started by taking a proposition as a target and then through the semantic units of WM it will construct appropriate branches to make a path. This process will be empowered through already available propositions in the LTM by facilitating partial structures. A key innovation in LISA was that it treats analogical mapping as a form of learning. LISA gradually learns the weights of mapping connections between appropriate elements that leads to a global solution without the need of massively parallel constraint satisfaction.

Due to the parallel nature of the model, and also the improvement of hardware, it has some scalability over a large data set (Hummel et al., 1997). Nevertheless, this analogy making model is based on mappings among propositions but not on domains; as a result of that LISA has some limitations over large domains on novel analogies, and over the size too (Keane et al., 1988).

6.3.6. Statistical driven Analogy Makings

CopyCat is a statistical driven emergent architecture which is neither symbolic nor connectionist (Hofstadter et al., 1995). The program’s macroscopic behaviour emerges from the interaction of a large number of low-level activities in which probabilistic decisions are made. Mental fluidity (slippages induced by pressures) is the concept that governs the CopyCat model’s functionality. CopyCat operates stochastically; therefore, it will not guarantee to produce the same answer on the same problem again. CopyCat solves letter-string analogies of the form: “ABC : ABD :: KJL : ? ” and gives probabilistically possible answers like LJK, KJJ, KJD, etc. (French, 2002). This model has three major components namely; Slipnet (LTM, that includes permanent platonic concepts), Workspace (like a working memory that consists with instances of types from the Slipnet), and Code-rack (agents who want to act on things in workspace reside here). Small teams of Code-rack agents explore different possibilities for structures, building based on what previous teams have constructed. This process leads to an emergent understanding of the analogy (Hofstadter et al., 1995). However, due to CopyCat providing a purely stochastic explanation and thus the factors contributing to the variability of problem solving are not clarified (Kokinov, 1994).

CopyCat lacks any episodic memory (long-term memory) and does not address the problem of accessing a source analogue from a large pool of past data.
Furthermore, ambiguity handling is a tricky part in letter-strings, which has to be justified through CopyCat to perform higher order analogy makings. Tabletop was a model that presented CopyCat’s architecture into a real world domain problem that deals with arrangements of objects on a table (French, 1995).

### 6.3.7. Associative Memory Based Reasoning

Associative Memory Based Reasoning (AMBR) is an analogy maker based on cognitive architecture called DUAL (Kokinov, 1994). In the DUAL architecture, which was inspired by Minsky’s book “The Society of Mind”, main functional elements are called DUAL agents who are very small entities but occur in large numbers (Kokinov et al., 2003). Main types of AMBR agents are Instance agent, Concept agent, Hypothesis agent, and Winner hypothesis (Kokinov et al., 2008). These agents collectively activate when required and do not use a central executive to control them, therefore through the interaction results will emerge (Petrov et al., 1998).

There are five phases identified in analogy making namely; retrieval, mapping, transfer, evaluation and learning, and in AMBR, these phases communicate with each other in parallel through DUAL agents. The learning phase is the unique feature of AMBR in comparison to all the other models that have been discussed so far. There are two mechanisms of retrieval in AMBR: automatic and strategic. Automatic retrieval is the process responsible for keeping the memory state of the reasoner in correspondence with the current context whilst strategic retrieval is used when the reasoner wants to reveal the relation between two concepts or situations at some stage of the reasoning process (Kokinov, 1994).

AMBR also uses structural, semantic, and pragmatic constraints in the mapping process. Main mechanisms in AMBR architecture are Spreading activation, Marker emission and passing, Structural correspondences, and Constraint satisfaction network (Kokinov et al., 2008). AMBR can be summarized as a hybrid-dynamic-emergent-cognitive model. While AMBR is both exciting and intuitive, AMBR’s ability to learn is quite limited. It is unable to identify the salient or statistically significant features of an episode, and it is unable to construct new representations of episodes or concepts (Dreisiger, 2008). AMBR has been improved to successor versions like AMBR2, AMBR3.

### 6.3.8. Overall discussion about Analogy Models

It is clear that almost all the latest achievements (Falkenhainer et al., 1986; Gentner et al., 1991; Hofstadter et al., 1995; Holyoak et al., 1989; Hummel et al., 1996; Keane et al., 1988; Kokinov, 1994; Thagard et al., 1990) have been inspired by various theories of psychology as the rationale behind those models. In this source to domain analogy making process, retrieval has obtained significant
attention and semantic similarities were the main concerns for this, whilst analogical mapping has been driven by structural similarities. Most classical models (SME, MAC/FAC, ARCS, and ACME) assumed sequential processing in a path in which first the retrieval process finds the base for analogy and then the mapping process builds the correspondences between the target and the retrieved base (Petrov et al., 1998).

Models like MAC/FAC and ACME contributed significant improvements to SME; nevertheless, computations required for such models are unrealistic from the perspective of the limitations of human working memory. Furthermore, in the ARCS and ACME models, the restrictions like allowing relations only with the same number of arguments to be mapped (n-ary restriction) have diluted the strength of those by comparing with the human process, which allows variable length mapping in real life. Due to the limitations such as time complexity noted in this sequential approach over a large data set and mapping restrictions; researchers’ focus moved to a parallel paradigm in which retrieval and mapping processes are asynchronous (Petrov et al., 1998).

LISA is an example for parallel architecture. LISA has lot of features that will push it relatively closer to a realistic model through features like: connectionist model, variable length mapping allowed, integrated memory retrieval and analogical mapping model. However, it is still far from the real brain mechanisms, and it can hardly deal with complex analogies.

Further research realizes that analogy making should not be restricted to be only parallel but useful to include multiple agent driven concepts. AMBR and CopyCat both have been inspired by this multi agent perspective. The main concern with CopyCat is that it has been developed for an abstract domain (letter strings), and certain generalization is required to use it as a general model. Furthermore, CopyCat does not have an episodic memory that stores old experiences. Nevertheless, there is no argument that these two models have acquired significant attention in analogy research due to the strong conceptual basis provide by them.

AMBR is different from all the other models, and it clearly shows its acceptance through practical applications too (Kokinov et al., 2008). However, the main concern with this approach is the confidence that we have in the DUAL architecture in terms of human cognition. Furthermore, these are inspired by the connectionist approach with neural network features, that is a purely mathematical, abstract model of the human brain. It is true that the power that can be derived from a connectionist model is fascinating, nevertheless practical findings have confirmed that the results that were generated are merely not for the problem that was wanted to be solved. In connectionist models, we accept the outputs as far as we are getting output relevant for the original question that we wanted to solve. Therefore, research on analogy making is still open.
6.4. Emergent Intelligence

In nature it is quite visible that some properties are not ordinarily associated with most of their physical entities. Atoms interact with one another, and hence generate emergent outcomes; for example two Hydrogen (H) atoms with one Oxygen (O) atom will result Water (H₂O), which merely contains hydrogen and oxygen, nevertheless properties of water cannot easily be reduced to properties of those two atoms. Furthermore, we know that water has three forms merely due to the dynamic states of the water molecules; namely ice, liquid, and steam. Each form has its own unique properties though each form is nothing but a set of water molecules.

Flame, is another good example to explain emergent phenomena, specifically in a time and space illusion. We know that fire is something that will emerge only when the required conditions are met (combustible substance, oxygen, heat, etc.). Nevertheless, if we ask where that flame was before those conditions were met, no good answer can be given.

Intelligence is also an emergent phenomenon. The difficulty with this suggestion on intelligence is that we have poorly understood it, and the criteria for intelligence as being emergent are very unclear. Nevertheless, according to the Buddhist philosophy, everything is considered to be as a conditioned phenomenon (Jayatilleke, 1978). When the conditions are met, appropriate entities (physical or non-physical) will emerge. Therefore, intelligence also will emerge when the appropriate conditions are met. The required criteria or ingredients for this emergent intelligence can be borrowed from cognition.

6.5. Cognition through Five Aggregates

Cognitive science is an interdisciplinary study of mind and intelligence in which representation and computation are considered as main challenges (Thagard, 2010). Representation can be achieved through predicate calculus, semantic net, ontology, etc. Ontology has gained success over all the other representation techniques due to its ability of connecting various forms of knowledge together.

Computation in cognitive modelling has been addressed through various approaches. A view on computation of cognition can be explored from a Buddhist perspective through a process called ‘Five Aggregates’. Five Aggregates are comprised of ‘Form’ (rūpa), ‘Sensation’ (vedanā), ‘Perception’ (saññā), ‘Mental formations’ (saṅkhāra), and ‘Consciousness’ (viññāṇa) (Piyadassi, 1972).

- **Form**: Form corresponds to material or physical factors which can be external or internal. The aggregate of form includes the five physical organs (eye, ear, nose, tongue, body, and mind), and the corresponding physical objects of the sense organs (sight, sound, smell, taste, tangible...
objects, and thoughts). This leads to acting as stimuli to the mind (through the senses). In other words form is the input to the mind.

- Sensation: Sensation is the basic acceptance of the input. Sensation will be experienced through seeing a form or a visible object, hearing a sound, smelling an odour, tasting a flavour, touching some tangible thing, cognizing a mental object (an idea or thought). Sensations are threefold: pleasant, unpleasant and neutral. Sensation bridges the internal sense organs with external sense objects. Furthermore, this deals with emotional dimensions too.

- Perception: The function of perception is to recognize objects, both physical and mental, from what we sense by seeing, hearing, smelling, tasting, feeling or thinking. It makes the assimilation of sensation information with ideas that pre-exist (a labelling process). This deals with conceptual dimensions.

- Mental formations: This describes a conditioned response to the object of experience. Through this we will develop thoughts. Furthermore, this is about the reasoning, which itself emerges by meeting the conditions stemming from already gather knowledge together with perceptions. According to Buddhism, there is no physical location to store the previous knowledge, yet emerged or recalled when meeting the conditions.

- Consciousness: Consciousness is mere awareness, or sensitivity to an object through other aggregates. Also this is nothing but the ready-state of the mind.

This is the main intuition in this research for analogy making. The actual process explained in Buddhist literature was not used as it is in this work, mainly due to the complexity and difficulties in being expressed in a model. Instead, an abstract model created out of five aggregates will be used for this purpose. Aggregates are dynamic processes, and conditioning through these factors enables us to perceive and understand the world, and the five aggregates work together to produce a mental being as per the Buddhist literature (Piyadassi, 1972). The abstract process were extracted from the five aggregates based literature (Piyadassi, 1972; Boisvert, 1995; Thera, 2008). Once the sense objects (sight, sound, smell, taste, tangible objects, and thoughts) meet with sense-organs (eye, ear, nose, tongue, body, and mind) that leads to a sensation that gives a feeling of either pleasant, unpleasant or neutral. The feeling that develops for the received inputs immediately leads to generate our perception. Through the perception, the external data will be recognised (or identified) and labelled (as this is a car, a house, a friend). Nevertheless, this aggregate can produce false recognitions which lead to wrong results (for example a rope may recognised as a snake). Therefore, proper
recognition is important for a proper understanding. Having labelled information of external input through a recognition process, this triggers the mental formation aggregate that provides a conditioned response to the recognised objects (or things). In Buddhism fifty two mental factors are considered for this process, but those will be not be addressed within the scope of this research. Mental formation is assumed to be something made of a combination of other things. Together with the generated feeling (through sensation), recognised information of the external (or internal) input (through perception), and the past knowledge/experience leads to form, fabricate, or construct a mental picture, that is conditioned to select an action or a thought. These conditioned phenomena include reasoning and with all background information and knowledge/experience it is decided what is to be done through comparing and contrasting. Finally, this leads to consciousness and development of a strong bond with other aggregates. Through this we are able to cycle again and again through the above steps and develop new thoughts and actions based on the external input received. Furthermore, consciousness drives each aggregate to shift its choices (for example to change the perception due to the internal conflicts is captured by the reasoning). Through this it is assumed that intelligence is emerging.

6.6. Approach and Design

This paper postulates that Analogy Making can be modelled as emergent intelligence from an abstract model that is created out of the Five Aggregates explained in Buddhist philosophy. In an analogy making process, a phenomena or a problem can be considered as a form and through sensation that will be mapped into a computer understandable form. This further contributes to relating it to existing ontology to get perceptions to identify the entities of it. That is further combined with concept formation techniques with the support of relations and having consciousness. In this approach multi-agent modelling is used to bridge the connection among aggregates and to facilitate intelligence as an emergent phenomenon. The detailed process will be as follows.

A given analogy problem will be considered as a ‘Form’ which is the actual input to the system. ‘Form’ will be either physical input or it can be a mental input which resulted from remaining processes. ‘Form’ needs to be mapped to a computer understandable form through ‘Sensation’. In this approach ‘Form’ to ‘Sensation’ plays the role of a transformation process. ‘Sensation’ agents take the given problem and transform it into its computer understandable form. The key concepts on this transformation task are abstraction and generalization due to the requirement of reduction of higher dimensionality and complexity. Computer understandable form may be represented in First Order Logic, Temporal Logic, Epistemic Logic, et cetera. Furthermore, by getting some heuristic information about the focus of the problem, the ‘Sensation’ process will contribute to guide the remaining processes.
In Buddhism, ‘Sensations’ can be threefold (pleasant, unpleasant and neutral), which directs our consciousness according to that feeling. Nevertheless, in this approach for analogy making it will be assumed that this agent will assign a level of difficulty (or complexity) in the given analogy problem: less, moderate and strong.

Perception primarily recognizes the entities, concepts, and attributes which are obtained from the ‘Sensation’ aggregate agent. In the human process, we use our memory for this purpose, which is considered as ontology in this analogy making model. According to Buddhism there is no physical location to store the previous knowledge, yet it emerges when meeting the necessary conditions. However, from a computational viewpoint, previous knowledge should be stored in ontology, and enable the emergence through conditioning among knowledge entities. In this regard an appropriate ontology has to be used for perception retrieval (selecting a wrong domain under the ontology leads to generate wrong perceptions).

A mental formation process is forming new imagination in the mind that may have not been previously experienced, or at least only partially formed. Imagination is considered to be more important than knowledge, as knowledge is limited. The main concept which is used in this process is “relations”. Through relations only we can perceive this world. Therefore, this process either establishes or fabricates possible relations as hypothesis by taking information from perceptions. In this task perception data will be used as the working memory. The domain identification for target is the most important task in this process, which will handle through ‘Mental formation’ aggregate. Domain specific knowledge is stored in the ontology (Long Term Memory) and based on the extracted relations from source, suitable list of domains will be short listed. This mainly occurs through the autonomy feature of agents. Agents are assigned to various domains; depending on the significance of adjacency and remoteness among domains the relevant agents will react. Through this process suitable domains will be short listed. Thereafter, conflict resolution techniques will be activated and isolate the highest potential domain(s). By retrieving the most positive mappings are allowed to make successful analogies. Therefore, this can strengthen them, together with some advanced techniques though those are not under the scope of the current work. Finally, agents share the findings into the working memory to proceed.

Consciousness is the ready-state of the mind. There is no arising of consciousness without conditions (Piyadassi, 1972). Buddhism has not defined what consciousness is, but it says that it is possible to recognize and identify it experientially. In a computational viewpoint this is a validation process that critically analyses the thoughts derived through mental formations. All the processes are activated appropriately based on concerns/questions/doubts from consciousness and will do the needful in a cyclic manner until a successful analogy will emerge as per Figure 1.
Fig. 1: Overview of the process of Five Aggregates

The significant difference of this proposed model over the other credited analogy models is that this is neither purely a search based nor connectionist driven, but powered through a cognitive model which has a philosophical basis in Buddhism. This model progresses through conditioning data or/and information that are available in each state in the model while cultivating intelligence. Figure 1 shows the abstract flow of the model; this is a cyclic process that takes the given analogy problem to relevant aggregates and ‘Consciousness’ will redirect the flow to strengthen quality of final analogy.

6.7. Geometric Analogy Solver

The above approach has been implemented for empirical validation in a Geometric Analogy domain, as shown in Figure 2. Geometric Analogies are less complicated and used as a testing domain often (Evans, 1964). Geometric Analogies are highly used in Intelligence Quotient (IQ) tests to assess human

Fig. 2: Multi Agent driven analogy making model with an abstract model created out of Five Aggregates.
intelligence, therefore through this it can measure the power of this approach in a comparable manner.

Java Agent DEvelopment Framework (JADE) was used with its inbuilt Agent Communication Language (ACL) to model the agent behaviour. As the knowledge source of agents, ontology was developed in the geometric domain. The quality of the model depends heavily on the adequacy of knowledge in the ontology; a light weight geometric ontology was used with adequate concepts, predicates, and facts. For the simplicity of the implementation of ‘Form’ and ‘Sensation’ has been combined to facilitate a direct input to the model in the form of First Order Logic (FOL), rather than having a transformation mechanism (actual transformation technique requires various image processing techniques and much more work).

Figure 3 shows an analogy problem in configuration form that was used for empirical validation. The problem was transformed into its FOL form as; above(0.A, 1.A), left-of(0.B, 1.B), no-change(0.A, 0.B), no-change(1.A, 1.B), above(0.C, 1.C), left-of (1.i, 0.i), no-change(0.C, 0.i), no-change(1.C, 1.i) et cetera. where each predicate is a Java object. Due to its expressive power first order logic through quantifiers (universal (∀) and existential (∃)) and the use of connectives {not (~), and (^), or (v), implies (→), if and only if (↔)} will be able to handle large number of domains effectively. The Perception agent will use these FOL data (which are Java objects) and starts of attaining beliefs like ‘0.A’ is a triangle which is a shape, ‘1.A’ is a square which is a special form of a rectangular shape, etc. These perceptions should include both specific and generic perceptions on the given problem for quality formations. Through this process, the model will construct its

![Fig. 3: Representations of the geometric analogy problem.](image-url)
Mental formation agents will induce possible mappings on identified relations from perceptions together with input information. These induced mappings are mainly inspired by transformations of simpler relations in geometric domain. Translation, reflection, dilation, and rotation are used as transformation techniques, and agents are individually working on these and trying to bridge the gap from source to target through transformation mappings. The Consciousness agent will validate these and do what is needed to improve the solution by getting the support from the other aggregates in a cyclic manner as shown in Figure 1. The Consciousness agent acts as a typical message space agent (as in Figure 2) in a multi-agent system. All the agents compile their findings into the message space (which has similarities to a Blackboard (Corkill, 1991), or Global Workspace (Baars, 1988, 1997)), and it acts as a bridging interface to enable collaboration among the agents, and coordination. Here intelligence emerges due to the overall interaction and communication among the agents through the message space. In a particular situation, the Consciousness agent will identify a particular thought as the answer which explains the analogy sufficiently. Nevertheless, in this model it is not possible to justify whether a given answer is the most suitable. Similarly if the same question is given to a human the answer will depend on his/her knowledge and reasoning abilities. The agent solution in this setup is assumed to be a rational choice.

6.8. Evaluation and Conclusion

The model was evaluated in comparative manner using a Turing Test approach; and Table 1 summarizes its findings. Ten geometric analogy problems were shared with the implemented model and the same analogies were given to 20 grade five students. Statistics shows ~70% percent success of the model, when compared to students. The selected analogy questions were mainly inspired by Evan’s questions which have been used for many other models too. Therefore, comparatively this model shows its strength among other models in geometric domain.

This paper discusses a different approach to handle analogies. The hypothesis was empirically validated in a comparative manner in the geometric analogy domain and obtained satisfactory results, while presenting a generic approach for analogy making. Nevertheless, this needs to be validated more extensively for complex analogies. Separation of Form and Sensation for the implementation phase will be considered as a separate research issue. It is noted that each process of this model can be empowered through methodologies discussed in Section 6.2 by forming a hybrid model but driven through Five Aggregates model.
Table 1: Experimental results (columns are for questions, rows are for students and the last row is for the model).

<table>
<thead>
<tr>
<th>Analogy Questions</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 1</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Student 2</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Student 3</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
</tr>
<tr>
<td>Student 4</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Student 5</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Student 6</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Student 7</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Student 8</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Student 9</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Student 10</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Student 11</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Student 12</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>°</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Student 13</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>°</td>
<td>✔</td>
<td>°</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Student 14</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Student 15</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Student 16</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Student 17</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>°</td>
<td>✔</td>
<td>°</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Student 18</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>°</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Student 19</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Student 20</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>MAS</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
</tr>
</tbody>
</table>

References


Evans T. G. (1964), A heuristic program to solve geometric-analogy problems. AFIPS 64 (Spring) Proceedings of the April, spring joint computer conference, pp. 21-23.


Chapter 7

Modelling the Role of Cognitive Metaphors in Joint Decision Making

Laila van Ments, Dilhan J. Thilakarathne, Jan Treur

Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.
e-mail: lailavanments@hotmail.com, d.j.thilakarathne@vu.nl, j.treur@vu.nl

Abstract: In this paper, a social agent model is presented for the influence of cognitive metaphors on joint decision making processes. The social agent model is based on mechanisms known from cognitive and social neuroscience and cognitive metaphor theory. The model was illustrated in particular for two types of metaphors that can affect joint decision making in different manners: a cooperative metaphor and a competitive metaphor. By a number of scenarios it was shown how the obtained social agent model can be used to simulate and analyze joint decision processes influenced by cognitive metaphor.

Keywords: metaphor, joint decision making, agent model.

---

1 This chapter was published as:
http://doi.org/10.1109/WI-IAT.2015.141

The names of the authors are ordered alphabetically reflecting the comparable contribution of each author.
7.1. Introduction & Motivation

Making decisions together with others is an essential part of human life, especially in social and professional (e.g., politics, trading, personal relationships) context. Choices in such a decision are always influenced by the others. This process is supported by an innate cognitive capability. Understanding and modeling the cognitive basis behind such joint decision-making is a non-trivial research challenge. Recent research from Social Neuroscience provides insight in some of the mechanisms underlying the different elements in joint decision-making processes; e.g., [1], [2].

Two core processes were identified as important to model processes of joint decision making: mirror neurons and internal simulation (see [3], [4]). Mirror neurons can prepare the body for a certain action or body change, and also activate upon observing somebody else who is performing or tending to perform this action or body change (e.g., [3], [5], [6]). Internal simulation is used as a means for prediction of the (expected) effects of a prepared action (e.g., [7]). Furthermore, internal simulation together with mirror neurons in an individual leads to the copying of processes that may take place in another individual. Hereby it facilitates empathic understanding associated to action choices of that individual, without actually executing those (e.g., [8], [9], [10]). An interplay of these processes among individuals can lead to the emergence of a joint decision. Additionally, ownership states also play an important role in decision-making process. An important function of an ownership state is that it mainly determines to what extent an individual attributes an action to himself or to another person. Ownership states (i.e., self and other ownership) are used together with prediction of the effects of a prepared action; they can influence whether the action is actually executed, and are the basis for acknowledging authorship of actions [11], [12]. Together, these concepts contribute to mutual empathic understanding between two persons, helping to form a well-founded joint decision. According to [3], a well-founded joint decision occurs when three requirements are met. Firstly, both agents have chosen the same option, secondly, both agents have a good feeling about it, and finally, both agents have empathic understanding of how the other feels about the chosen option.

In addition to the above processes and aspects, other aspects may still influence a joint decision making process. One perspective to explore this is through cognitive metaphor [13]-[19]. Originally, the theory was that a metaphor only occurs in written or spoken speech. However, according to George Lakoff and Marc Johnson [20] metaphors regularly play an important role in our conceptualization and communication and contribute to a mental image of a current situation. The way we understand many phenomena in our daily lives is metaphorical: one certain mental domain is understood in terms of another mental phenomenon. According to [21], metaphors are not only a figure of speech, but a mode of thought, that is
systematically mapped in our brains. The function of metaphor is to better understand certain concepts and it is an inevitable part of human thought and reasoning [21]. The cognitive theory on metaphors can be difficult to comprehend because metaphors are deeply and unconsciously engraved in our brain, making them so mundane to us that we do not notice them ourselves. Still, metaphors structure the way we think, how we see the world, and also the way we make decisions together with others. It has been found that the process of human thought through metaphorical associations can unconsciously be affected by bodily changes (see [22]).

In this paper the influence of cognitive metaphors on joint decision making will be examined by combining the concept of a joint decision making process and cognitive metaphor together in a computational social agent model. The computational social agent model for joint decision making presented in [3] is taken as a point of departure and extended by incorporating a model for metaphors.

Many kinds of metaphors can influence joint decision processes. The introduced model in this paper indicates in a generic manner how they can be incorporated in a model of joint decision making. In addition, it will be shown more specifically how influences of cooperative and competitive metaphors can be modelled in a detailed manner. This will provide a clear illustration of the influence cognitive metaphors may have in a joint decision making process.

The model proposed in this paper can have many possible uses. For example, it could be helpful in the design of human-like virtual agents for simulation-based training, like a virtual agent helping a human in training (group) decision tasks. Another possible use for a more specific task is a virtual agent that helps partners to learn how to make solid joint decisions, or that helps a business to come to a most valuable deal with another business. A model like this is useful in complex simulations of socio-technical systems (e.g., the aviation domain) based on computational models of human behaviour for more realistic results. In this paper, first in Section II some of the core concepts used are briefly reviewed. In Section III, the computational social agent model is presented. In Section IV, some of the explored simulation scenarios are discussed. Section V is a discussion and will bring forward some ideas for future research.

7.2. Core Concepts: Mirroring, Internal Simulation, Ownership, and Metaphors

The agent model presented here includes a number of cognitive states and processes. The theoretical basis of these states and processes is discussed in the current section.
7.2.1. Mirror neurons

Mirror neurons are fundamental to joint decision making. Mirror neurons are motor neurons that fire when an action is (to be) executed by a subject, but also when the subject observes somebody else performing that action. Observing an action activates the same neural mechanism as preparing for execution of said action [23]. This means that when an action is executed by someone else, this is not just perceived and represented in a sensory manner, also a motor representation occurs in the observers’ mind. Mirror neurons were originally found in monkeys [24], but later studies have proven the existence of a similar mechanism in human beings [5], [25], [26]. Gallese [23] explains that the mirror neuron areas in one’s brain are responsible for the processes of action execution, action perception, imitation and imagination, with neural links to motor effectors. When an action is executed or imitated, this leads to the excitation of the muscles concerned with that action. When an action is only observed or imagined, the excitation of the muscles does not take place.

7.2.2. Process of internal simulation

Internal simulation is another crucial concept in a joint decision making process. This works in combination with mirror neurons. The mirror neuron function makes that a preparation state is activated upon observing an action. As a next step, internal simulation generates a prediction of the (expected) effects of such a prepared action [7].

Emotions and feelings are important elements in human cognition. William James proposed that, after a person receives an input, as a response the body prepares for and executes bodily changes (referred to as body-loop) and only then feels an emotion. [27]. In short:

\[
\text{sensory representation} \rightarrow \text{preparation for bodily changes} \rightarrow \text{expressed bodily changes} \rightarrow \text{emotion felt} = \text{based on sensory representation of bodily changes}
\]

Damasio [8] introduced another hypothesis: the as-if body loop. The as-if body loop makes it possible that actual bodily changes in the emotion generation are bypassed by internally simulating the body changes. A person is faced with an input, this input leads to a preparation for bodily changes, and as a form of internal simulation, this leads to a (sensory) representation of a changed body state, causing the emotion that is felt, without actually executing the bodily changes. In addition, Damasio adds that the felt emotion and the preparation for bodily changes mutually affect each other, leading to a cycle. In short:
sensory representation $\rightarrow$ preparation for bodily changes $=$ emotional response $\rightarrow$ emotion felt $=$ based on sensory representation of (simulated) bodily changes $\rightarrow$ preparation for bodily changes $=$ emotional response

In combination, mirror neurons and as-if body loops can make feelings and actions of two persons converge. For example, person A gets sensory input of an action that person B (tends to) executes, and of person B’s associated emotion. Then the mirror neurons of person A lead to a preparation state in person A of the action that person B executes and the associated emotion. This, through the as-if body loop, will lead to person A having feelings and preparations that correspond to the action that person B executes and to B’s associated emotion. This mechanism explains how persons affect each other’s decisions and feelings so that convergence can be facilitated [3], [4], [6].

7.2.3. Ownership

Differentiating between the actions that are caused by oneself and actions that are caused by others is a vital concept of joint decision making and a requirement for establishing social communication and appropriate interactions. This system of self-other differentiation is important for functions as understanding the meaning of an action and interpreting the meaning of the responsible agent [12], [28], [29]. Another element, put forward by Moore and Haggard in [11], is the distinction between action ownership based on prediction (prior to execution), and action ownership based on inference after execution of the action (in retrospect). Only when prior to executing an action in the internal simulation process a person attributes the simulated action to himself and predicts the action to have a good outcome, this can result in actual execution of the action. Therefore, prior ownership states play an important role in controlling the actual execution of actions (go/no-go decisions, vetoing). After the execution, the person responsible for executing the action can acknowledge in retrospect the ownership of the action. This acknowledgement is necessary to enable communication of feelings and understanding about an action between people (see [12]).

7.2.4. Empathic response

In [30], p. 435 the following criteria for empathy are expressed for a person (S) having a state of empathy for another person (B): (1) presence of an affective state in the person, (2) isomorphism of the person’s own and the other person’s affective state, (3) elicitation of the person’s affective state upon observation or imagination of the other person’s affective state, and (4) knowledge of the person that the other person’s affective state is the source of the person’s own affective state. Assuming true, faithful bodily (nonverbal) and verbal expression, the following reformulation
can be made to obtain criteria for an empathic response to another person. If the prepared body state is actually expressed by $S$, so that the other agent $B$ can notice it, then this contributes an *empathic nonverbal response* of $S$ to $B$, whereas communication of $S$ of the emotion to $B$ (i.e., $S$ is communicating that $B$ has this emotion) is considered an *empathic verbal response*. The bodily expression of an observed emotion together with such a communication to $B$ occurring at the same time is considered a *full empathic response* of $S$ to $B$.

### 7.2.5. Cognitive metaphors

A person may encounter many unknown situations; the brain somehow has to make sense of them. According to cognitive metaphor theory, our brain maps knowledge of known concepts onto new ones to comprehend new situations [31]–[33]. This is also referred to as analogy making: a mapping between two domains, called the source (or base) and the target (or topic) [31], [32], based on a number of features or characteristics the base and the topic have in common. Consider for example as a metaphor the sentence *‘That woman is poison’*. Literally, this does not make sense; a human being is not a venomous object. However, this sentence can be recognized as a cognitive metaphor, with *‘woman’* as the topic and *‘poison’* as the source. This might lead to conceiving said woman as something that kills, injures, or impairs an organism and is something destructive or harmful.

As described by El Refaie [34], metaphors can change the way we think about a situation, as constant repetition of particular metaphors will lead to our unconscious acceptance of that particular metaphor as a normal way of seeing that situation. Thereby, a metaphor subconsciously constructs how we perceive concepts (see [22]). Furthermore, several studies have shown that our actions are subconsciously influenced by the automated activation of motives [35], [36]. In particular, the concepts and motives playing a role in joint decision making process, including all underlying processes, are influenced by our metaphorical image of the situation. Therefore, cognitive metaphors strongly affect the way humans make decisions with each other. In this paper, the influence of cognitive metaphor on the joint decision making process will be examined and illustrated, specifically for two categories of metaphors: a cooperative metaphor and a competitive metaphor. For example, if a person uses the metaphor *‘war’* to make a decision, he will unconsciously experience the decision making process as a form of war, attacking the opponent and defending his own points. This will lead to a competitive mind set and an outcome with one winner and one loser. However, if a person uses a friendlier metaphor for the decision process, for instance *‘art’*, this will lead to a cooperative mind set. If a person uses this mindset in the joint decision making process, he or she will aim at creating something together with the partner, leading to a joint outcome.
7.3. A Computational Dynamic Social Agent Model for Joint Decision Making with Metaphors

A neurologically inspired social agent model is presented that integrates the role of metaphors in joint decision making. It adopts elements of previously developed models, in particular models on joint decision making processes [3] and ownership [12]. Based on these elements and the knowledge about cognitive metaphors as discussed, an integrative social agent model was created that focuses on the influence of cognitive metaphor in joint decision making processes. For a graphical overview of the model, see Fig. 1 and for its abbreviations, see Table I. The graphical overview incorporates several states (circles) and their dynamics as processes (arrows). This model represents the cognitive behaviour of one agent (Self: S) and it observes the actions and emotions of another agent (agent B) in the process of joint decision making.

7.3.1. Basic elements in the joint decision making model

The model uses five world states (WS) as inputs: stimulus $s$: WS($s$), action $a$ of agent B: WS($B, a$), effect $e$ of self action $a$: WS($e$), feeling $b$ associated to action effect $e$ of Agent B: WS($B, b$), and feeling $b$ associated to self action effect $e$: WS($b$). In Fig. 1 it has used subscript letter ‘i’ for features: $a$, $e$, and $b$; that indicates the $i^{th}$

![Fig. 1: A graphical overview of the social agent model](image-url)
instance of that feature. Therefore, through this model it is possible to have multiple action options through a single stimulus. The input world states $\text{WS}(s)$, $\text{WS}(B, a_i)$, $\text{WS}(e_i)$, $\text{WS}(B, b_i)$, and $\text{WS}(b_i)$ lead to sensor states $\text{SS}(s)$, $\text{SS}(B, a_i)$, $\text{SS}(e_i)$, $\text{SS}(B, b_i)$, and $\text{SS}(b_i)$, and subsequently to sensory representation states $\text{SR}(s)$, $\text{SR}(B, a_i)$, $\text{SR}(e_i)$, $\text{SR}(B, b_i)$, and $\text{SR}(b_i)$, respectively.

From the dynamic perspective, at a given point in time two or more agents receive/observe a stimulus $s$. This leads to a causal chain $\text{WS}(s)$ to $\text{SS}(s)$ to $\text{SR}(s)$. Having sensory representation of the stimulus $s$, each agent individually prepares for possible options through preparation states $\text{PS}(a_i)$ for executing action $a_i$; these options may correspond to the habitual responses upon the stimulus. Having represented $\text{PS}(a_i)$ states a decision or selection process is needed to decide which option to select. For this, the internal simulation process is used (following Damasio): without executing the action options through the body loop, each option is internally simulated by the agent to determine and feel the predicted effect. Therefore, internally the agent develops sensory representations $\text{SR}(e_i)$ of the effect of actions $a_i$, which provides a basis for evaluation of each option. In the model this process is called the prediction loop. Furthermore, to express mutual exclusiveness between some options, if relevant in a scenario, a suppressive link from $\text{PS}(a_i)$ to itself is used. More specifically, if (and only if) two action options $a_i$ and $a_j$ with $j \neq i$ cannot occur together, then the $\text{PS}(a_i)$ state suppresses $\text{PA}(a_j)$ and conversely. This is shown as a dotted looped red arrow in Fig. 1. This behaviour is in line with the explanation for lateral inhibition in [37] and contributes to further strengthen the action selection process.

<table>
<thead>
<tr>
<th>Table I. Nomenclature for Fig. 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WS(W)</strong></td>
</tr>
<tr>
<td><strong>SS(W)</strong></td>
</tr>
<tr>
<td><strong>SR(W)</strong></td>
</tr>
<tr>
<td><strong>PS(Z)</strong></td>
</tr>
<tr>
<td>$Z \in {a, b}$ for action $a$ and effect $e$</td>
</tr>
<tr>
<td><strong>OS(X,s,a,e)</strong></td>
</tr>
<tr>
<td>$X \in {S, B}$</td>
</tr>
<tr>
<td><strong>OS(X,e,b)</strong></td>
</tr>
<tr>
<td>$X \in {S, B}$</td>
</tr>
<tr>
<td><strong>Met(Y)</strong></td>
</tr>
<tr>
<td>$Y \in {\text{coo, com}}$</td>
</tr>
<tr>
<td><strong>ES(Z)</strong></td>
</tr>
<tr>
<td>$Z \in {a, b}$</td>
</tr>
<tr>
<td><strong>EC(X,s,a,e)</strong></td>
</tr>
<tr>
<td>$X \in {S, B}$</td>
</tr>
<tr>
<td><strong>EC(X,e,b)</strong></td>
</tr>
<tr>
<td>$X \in {S, B}$</td>
</tr>
</tbody>
</table>
This model provides a selection of the option that has the strongest valuated feeling. Parallel to this, the agent develops a self-ownership state \( \text{OS}(S, s, a_i, e_i) \) for its action selection. This self-ownership strengthens the execution of that option and therefore together with having a strong predicted feeling the agent may move to action execution. Note that for the sake of simplicity this model does not include the differentiation of prior and retrospective states; for more information on this distinction, see [12]. While this action selection process is developing, the agent starts to execute basic indications of its (to be) selected option \( a_i \) through partial activation of its execution state \( \text{ES}(a_i) \). As this is in the context of joint decision making, the preferred signs of choices of each agent will be observed by the other agents (for the simplicity of explanation only two agents are considered). Therefore, the agent \( S \) observes that the other agent \( B \) tends to perform action \( a_i \) through its observed world state \( \text{WS}(B, a_i) \), leading to a sensory representation state \( \text{SR}(B, a_i) \). At this point of time the mirror neurons are used in this model. It is assumed that through the relevant mirror neuron function of preparation state \( \text{PS}(a_i) \), sensory representation state \( \text{SR}(B, a_i) \) affects \( \text{PS}(a_i) \). In this way observing the other agent \( B \) affects agent \( S \)’s corresponding states and preparations, leading to the feelings and decisions of both agents to be tuned to each other, as discussed in the previous section. Moreover, the agents differentiate the self and other (agent \( B \)’s) ownership and start to develop other-ownership states \( \text{OS}(B, s, a_i, e_i) \). Furthermore any ownership state \( \text{OS}(X, s, a_r, e_i) \) suppresses \( \text{SR}(e_i) \) after executing action \( a_i \). This is important for the separation of effects of action prediction and execution as highlighted in [11]. Due to this it is expected to have a dip in the sensory representation and feeling in-between predictive representation and inferential representation [38].

In this model, it is assumed that an agent will not perform an action spontaneously but starts to slowly provide signs of execution and develop or suppress it over time more and more strongly. In line with the agent’s initial preparation of action \( a_i \), it will add activation to \( \text{SR}(e_i) \). This will lead to emotions associated to the predicted effects of action \( a_i \): the agent prepares for expressing emotions for effect representation \( \text{SR}(e_i) \) through \( \text{PS}(b_i) \). Each emotion is evaluated through the process of internal simulation (by the as-if body loop in Fig. 1) and the agent experiences its associated feeling (without executing it) and in parallel develops the self-ownership of the emotion indicated by body state \( b_i \) and effect \( e_i \): \( \text{OS}(S, e_i, b_i) \). For mutually exclusive options, the state \( \text{PS}(b_i) \) also includes a self suppressive link as explained for \( \text{PS}(a_i) \).

Similar to the action \( a_i \), agents start to share the signs of their emotion through execution state: \( \text{ES}(b_i) \). As the same process is developing inside the other agent \( B \), agent \( S \) can see the emotions of agent \( B \) through \( \text{WS}(B, b_i) \) and represent this by \( \text{SR}(B, b_i) \). Through the mirror neuron function it also effects on \( \text{PS}(b_i) \) and leads to
develop OS($B, e_i, b_i$) (as explained for SR($B, a_i$)). Furthermore the ownership state OS($X, e_i, b_i$) also suppresses SR($b_i$) after executing $b_i$ [11] as explained for the OS($X, s, a_i, e_i$).

### 7.3.2. The integration of metaphors in the model

In a generic manner there are two sides for the (functional) role that characterizes a metaphor within the causal chains of mental processes: (1) how is it affected by certain states, and (2) how does it affect other states and processes.

Side (1) of a characterization of a metaphor specifies to which situations it applies. Through this it is determined in which situations a given metaphor becomes activated. This works according to a process of analogy making (see [31], [32]). This is modeled by considering a number of characteristics or features of a situation and match this to characteristics or features of the metaphor. If a match occurs, the metaphor will become active, else it will remain inactive. So, any given metaphor is characterized by such a set of characteristics or features. The strength of the matching process also involves personal characteristics of the agent. For a given situation one agent may match it strongly to one specific metaphor, whereas another agent may match it more strongly to an alternative metaphor for the same situation. In this way individual differences occur. In general the characteristics of the situation are internally represented by a set of sensory representations SR($w_1$), …SR($w_n$), for example, formed on the basis of observations SS($w_1$), …, SS($w_n$) of the situation. For a specific agent the matching of these characteristics to a metaphor state Met(m) is modeled by the agent-specific strengths of the connections from these sensory representation states SR($w_1$), …SR($w_n$) to the metaphor state Met(m). For the case of the specific metaphors relevant for joint decision making the relevant characteristics in the first place include the fact that there are other agents involved, and secondly, that there are different action options to consider.

Once a metaphor has become active, it affects other states and processes. This is the second part (2) of the characterization of a specific metaphor. For a given metaphor, this is modeled by specifying connections with certain (agent-specific) weights from the metaphor state to other states. For the case of the specific metaphors relevant for joint decision making such connections are to the states relevant in the joint decision making process. In this case a metaphor state of agent S influences the ownership states for action options and feelings. In this way, through the ownership states, the metaphor state has influence on whether an action option is chosen or not.

The specific metaphors used as illustration in this paper are the cooperative metaphor and the competitive metaphor. Both metaphors share as a characteristic that they only apply when another person is present and play a role in some form of
joint decision making in which different options are considered. So, to specify to which situations these metaphors apply (1) connections from sensory representation states $SR(B, ..)$ are used, in particular:

(1) Sensory representation states $SR(B, a_i)$ and $SR(B, b_i)$ activate cooperative and competitive metaphor states $Met(coo)$ and $Met(com)$

The strengths of these connections can be different between different agents, and also different between the cooperative and the competitive metaphor, thus also expressing personal characteristics of an agent.

The way in which the cooperative and competitive metaphor states have effects (2) on the decision making is as follows. These effects are modeled by connections from the metaphor states to ownership states $OS(X, ..)$ in such a way that:

(2a) A cooperative metaphor state $Met(coo)$ suppresses self-ownership states $OS(S, ..)$

(2b) A competitive metaphor state $Met(com)$ strengthens self-ownership states $OS(S, ..)$

By these effects an agent tends more to keep the own preferred option using a competitive metaphor and less using a cooperative metaphor.

Due to the processes as explained the agent will (strongly) execute an action $a_i$, and emotion $b_i$ through the body loop. For the execution of $ES(a_i)$ and $ES(b_i)$ only self-ownership states will have effect, in addition to the preparation states. Finally the agent acknowledges authority for the executed action and communicates about the effects and emotions of the executed action through $EC(X, s, a_i, e_i)$ and $EC(X, e_i, b_i)$.

7.3.3. Dynamics of the model

Having a cognitive model, it is important to translate it into a computational format to obtain simulation results. For this purpose, the model was compiled as proposed in [39] to simulate scenarios. Connections between state properties (the arrows in Fig. 1) have weights $\omega_k$, as indicated in Table II. In this table a weight $\omega_k$ has a value between -1 and +1 and may depend on the specific stimulus $s$, action $a_i$, effect $e_i$ and/or emotion $b_i$ involved. By varying these connection strengths, different possibilities for the repertoire offered by the model can be realized and can be aligned with considered scenarios and personal characteristics of agents. Usually weights are assumed to be nonnegative, except for inhibiting connections. The dynamics of the model depends on the values of each of these weights (together with other parameters).

Local Properties in LEADSTO (see [40]) specify the update dynamics of the activation value of the ‘to state’ based on the activation levels of the ‘from states’.
For the dynamics of each local property a LEADSTO formalisation was made\(^2\). Each state includes an additional parameter called speed factor \(\gamma_i\), indicating the speed by which an activation level is updated upon received input from other states to the state ‘\(i\)’. Two different speed values are used as fast and slow: fast value is for internal states and slow value is for external states (the states inside the dotted black block are considered as internal states). Activation of a state depends on multiple other states that are directly attached to it. Therefore incoming activation levels from other states are combined to some aggregated input and affect the activation level according to a differential equation as in (1) below. For each state a continuous logistic threshold function is used as in equation (2), where \(\sigma\) is the steepness, and \(\tau\) the threshold value. When the aggregated input is negative (3) is used. To achieve the temporal behaviour of each state as a dynamical system, a difference equation is used in the form of equation (4) (where \(\Delta t\) is the time step size).

\[
\frac{dy_i}{dt} = \gamma_i \left( f \left( \sum_j \omega_{ij} y_j \right) - y_i \right) \quad (1)
\]

\[
f(X) = th(\sigma, \tau, X) = \left( \frac{1}{1 + e^{-\sigma(X-\tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{\sigma\tau}) \text{ when } X > 0 \quad (2)
\]

\(^2\) http://www.few.vu.nl/~dte220/IAT2015_LEADSTO.pdf

For the dynamics of each local property a LEADSTO formalisation was made\(^2\). Each state includes an additional parameter called speed factor \(\gamma_i\), indicating the speed by which an activation level is updated upon received input from other states to the state ‘\(i\)’. Two different speed values are used as fast and slow: fast value is for internal states and slow value is for external states (the states inside the dotted black block are considered as internal states). Activation of a state depends on multiple other states that are directly attached to it. Therefore incoming activation levels from other states are combined to some aggregated input and affect the activation level according to a differential equation as in (1) below. For each state a continuous logistic threshold function is used as in equation (2), where \(\sigma\) is the steepness, and \(\tau\) the threshold value. When the aggregated input is negative (3) is used. To achieve the temporal behaviour of each state as a dynamical system, a difference equation is used in the form of equation (4) (where \(\Delta t\) is the time step size).

\[
\frac{dy_i}{dt} = \gamma_i \left( f \left( \sum_j \omega_{ij} y_j \right) - y_i \right) \quad (1)
\]

\[
f(X) = th(\sigma, \tau, X) = \left( \frac{1}{1 + e^{-\sigma(X-\tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{\sigma\tau}) \text{ when } X > 0 \quad (2)
\]

\(^2\) http://www.few.vu.nl/~dte220/IAT2015_LEADSTO.pdf
\[ f(X) = 0 \text{ when } X \leq 0 \]  
\[ y_i(t + \Delta t) = y_i(t) + \gamma_i \left[ \sum_{j \in s(i)} \omega_{ij} y_j \right] \Delta t \]

7.4. Simulation of Example Scenarios

The model presented in Section III was validated by simulating a large number of example scenarios based on literature discussed in Section II. Simulations have been generated using an implementation that was developed in Java using additional technologies to handle the input in XML format and to visualize the results. In each example scenario, some of the parameters of the model are varied, and the influence on the outcomes are examined. Each scenario consists of two agents, A and B, two action options, a1 and a2, and body state b1 and b2. Agent A always starts with a preference for action a1 and body state b1, agent B always starts with a preference for action a2 and body state b2. In this setup, there are three configurations possible regarding the metaphor state of both agents: both agents can be in a co-operative metaphor state, both agents can be in a competitive metaphor state, or one agent can be in a cooperative metaphor state while the other agent is in a competitive state. In the simulations, it will become clear if and how the metaphor states (together with the other discussed processes), have an influence on the outcome of the joint decision making process.

In the first scenario only the metaphor states are varied (i.e., link SR(B,X) to Met(Y)). The second scenario examines the influence of the mirroring process on the joint decision process (i.e., link SR(B,X) to PS(X)). In the third scenario the prediction loop is addressed (i.e., link PS(a_i) to SR(e_i)). Then, in the fourth scenario, the as-if body loop is varied (i.e., link PS(b_i) to SR(b_i)). Finally, in the fifth scenario, the ownership states are considered (i.e., link Met(Y) to OS(...)). Each of these scenarios is examined under three different values, namely low, medium and high, to validate the behaviour in joint decision making. This provides useful information to understand the model behaviour in different conditions or situations which is an important factor in application domains. Having realistic results for these scenarios confirms the capacity and adaptability of this model in joint decision making related simulation experiments.

In all of the scenario’s step size was taken 0.25, the slow speed factor was taken 0.5, fast speed factor was taken 0.8, and the total timeslot was 200. For each scenario different parameters are used, which are defined in Table III. Complete parameter values used for: both agents in cooperative metaphor state, both agents in competitive metaphor state, and one agent in cooperative metaphor state while the
other agent in cooperative metaphor state; can be found at\(^3\). In the Table III, only the specific changes relative to these files are given.

### 7.4.1. Varying the metaphor states

This scenario examines the influence of the metaphor state on the joint decision process with the model in the natural state, meaning that no processes in the model besides the metaphor state are varied. The following values were chosen for the connection strengths \(SR(B,\{a,b\}) \rightarrow Met(Y)\): 0.5, 0.8, 1.0. Under these three ranges, simulations should clearly show the influence of \(Met(Y)\) on joint decision. When the link value is low (i.e., 0.5), the agent should develop a relatively low \(Met(Y)\) state and therefore affects of joint decision should be diluted (or weakened), whereas when this link has a high value (i.e., 1.0) this provides the strongest effect from \(Met(Y)\) states and therefore either both agents select one choice (if metaphor states are cooperative) with stronger strength or each executes its own preference (if metaphor states are competitive) more strongly. The original simulation graphs for Scenario 1, the cooperative configuration with a link strength \(SR(B, X) \rightarrow Met(Y)\) (with \(X \in \{a,b\}\)) of 1, can be found in Fig. 2. The upper graph shows agent A, the lower graph shows B. Fig. 2 only includes the ownership, preparation, metaphor and execution states. More detailed figures with all the states can be found in an additional document [insert reference]. At time point 0 both agents are triggered by a stimulus s, and around time point 3 both agent start to prepare for action 1 and action 2. At time point 10, as bodily reactions, world states develop for each agent by executing the actions with their strengths, which, leads to observation by the other agent through its sensor states and finally sensory representations develop of this process (\(SR_B.a_i\)) at time point 20. As a result, around time point 23 the cooperative metaphor states are generated. After that, around time point 33, the other ownership states become stronger and the self ownership states are suppressed. At time point 85, the activation of the sensory representation of the other agents executing action 2 becomes stronger and the activation of the sensory representation of the other agents executing action 1 becomes weaker, leading to the ownership states of action 2 to become more activated and the ownership states of action 1 to become less activated at time point 85. This process keeps going, until finally around time point 130 the execution state of both agents of action 2 has become very high and the execution state of action 1 has become almost 0. At this time point also the communication about the execution of action 1 takes place and a solid joint decision has been formed. The process took some time because agent A had to switch preferences due to mirroring processes. The process of generating

\(^3\) [http://www.few.vu.nl/~dte220/IAT2015_xmlFiles.zip](http://www.few.vu.nl/~dte220/IAT2015_xmlFiles.zip)
feeling 2 and suppressing feeling 1 in this scenario is the same as the process for generating action 2 and suppressing action 1.

7.4.2. Adjusting the mirroring process

In Scenario 2, the process of mirroring between the agents is varied in the model. The process of mirroring can make actions and feelings of two people converge. If this process does not work normally, this can have impact on the joint decision making process. For instance, people with autism are sometimes assumed to suffer from impaired mirror neurons [41], leading to different behaviour in the joint

![Simulation graphs for Agent A & B for metaphor states, ownership states, preparation states, and execution states](http://www.few.vu.nl/~dte220/IAT2015-Fig2.pdf)
decision making process. Here the strength of the link \( SR(B, X) \rightarrow PS(X) \) with \( X \in \{a,b\} \) that represents the mirroring process, is varied (for each metaphor state configuration): 0.5, 0.7, 1.0.

### 7.4.3. Adjusting the prediction loop

In this scenario the prediction connection \( PS(a) \rightarrow SR(e) \) between the preparation state of an action and the sensory representation of the effect of that action is varied. For example persons with schizophrenia are assumed to suffer from impaired prediction [6], [42]. For some cases involving variation in these links, see Table III.

### 7.4.4. Adjusting the as-if body loop

In Scenario 4 the as-if body loop, the connection between the preparation state of a body state or feeling and the sensory representation of that body state or feeling is adjusted. When a person does not have a normally functioning as-if body loop, the internal simulation process as discussed in Section II does not work appropriately. This way, the emotional effect of an action cannot be predicted properly, which can have an impact on the joint decision making process. For some cases involving variation in these links, see Table III.

### 7.4.5. Adjusting the effect on ownership state

In the last scenario, the connection strength between the metaphor states and the ownership state is adjusted. This way, the strength of the influence of metaphor on the process is varied. The normal connection strengths of \( Met(m) \rightarrow OS(X) \) are presented in Table III. For each adapted strength, influence of several metaphor states and strengths are compared.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Nr</th>
<th>Agent</th>
<th>Metaphor configuration</th>
<th>Connection weights used</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted metaphor ( SR(B,X) \rightarrow Met(m) ) ( X \in {a,b} )</td>
<td>1a</td>
<td>A</td>
<td>Comp 1</td>
<td>Comp 1</td>
<td>Both agents execute their own preferred action and feeling and no mutual empathic understanding exists. No solid joint decision has been formed.</td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>A</td>
<td>Coop 1</td>
<td></td>
<td>Both agents execute action ( a_2 ) and body state ( b_2 ). Mutual empathic understanding exists, a solid joint decision has been formed. The process took longer than in the competitive scenario, because it takes time for agent A to switch his action preference from ( a_1 ) to ( a_2 ).</td>
</tr>
<tr>
<td>Adjusted mirroring ( SR(B,X) \rightarrow PS(X) )</td>
<td>2a</td>
<td>A</td>
<td>Coop 1</td>
<td></td>
<td>Here, finally both agents execute ( a_2 ) and ( b_2 ). Mutual empathic understanding exists, and so does a solid joint decision. This scenario shows more activation of ES of ( a_2 ) and ( b_2 ) than when the strength of the mirroring process (i.e., the link from ( SR(B,a) ) to ( PS(a) )) is weaker. Even though</td>
</tr>
<tr>
<td></td>
<td>2b</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Tabel III**: Summary of Simulation Results
| \(X \in \{a,b\}\) | 2b | A | Comp 0.5 | B | Comp 0.5 | Both agents are competitive and have an impaired mirroring process, meaning that they are less influenced by the actions of the other. The scenario has the same outcome as ‘1a’; however, for both agents their preferred action is executed more strongly and the other action has even less activation. |
| Adjusted prediction link | 3a | A | Coop 0.3,0.2 | B | Coop 0.3,0.6 | Both agents execute action \(a_2\) and body state \(b_2\). Mutual empathic understanding exists, a solid joint decision has been formed. In comparison to this configuration with stronger prediction link, there is less gap between the executed actions, e.g. the difference in strength in execution of \(a_1\) and \(b_1\) and \(a_2\) and \(b_2\). Also the ES states are less steep with a weaker prediction link. |
| \(PS(a) \rightarrow SR(e)\) | 3b | A | Coop 0.3,0.2 | B | Coop 0.3,0.6 | Both agents execute action \(a_2\) and body state \(b_2\). Mutual empathic understanding exists, a solid joint decision has been formed. For agent A, the lower prediction loop leads to less execution of all actions. For agent B, the lower prediction strength leads to a more competitive effect: \(a_2\) and \(b_2\) execute stronger, \(a_1\) and \(b_1\) execute less strong. |
| Adjusted as-if body loop | 4a | A | Coop 0.4,0.3 | B | Coop 0.3,0.6 | Both agents executed their own preferred action and feeling and no mutual empathic understanding exists, no solid joint decision has been formed. Compared to a stronger as-if body loop, the agents act more competitive, e.g. they execute their own preferred action more and the other action less strongly than with a stronger as-if body loop. |
| \(PS(b) \rightarrow SR(b)\) | 4b | A | Coop 0.5,0.3 | B | Coop 0.3,0.6 | Both agents execute action \(a_2\) and body state \(b_2\). Mutual empathic understanding exists, a solid joint decision has been formed. Preparation states, execution states en communication states are a lot less steep, making the decision making process slower. |
| Adjusted metaphor effect | 5a | A | Coop 0.4,-0.4; 0.4,-0.4; -0.2,0.2; -0.2,0.2; 0.4,-0.4; -0.2,0.2; -0.2,0.2 | B | Coop 0.4,-0.4; 0.4,-0.4; -0.2,0.2; -0.2,0.2; 0.4,-0.4; -0.2,0.2; -0.2,0.2 | Both agents execute action \(a_2\) and body state \(b_2\). Mutual empathic understanding exists, a solid joint decision has been formed. Comparing this scenario with a stronger connection between metaphor and ownership: for both agents, all other ownership states develop less strongly, self ownership states develop stronger due to less influence of cooperative metaphor. Due to this, it takes longer before both agents’ actions and feelings converge, than with a normal ownership state; the ES, EC graphs are less steep. |
| \(Met(y) \rightarrow OS(X,s,a,e)/\OS(X,e,b)/\OS(B,S)\) | 5b | A | Coop 0.8,-0.9; 0.8,-0.9; -0.5,0.5; -0.5,0.5; 0.8,-0.9; -0.5,0.5; -0.5,0.5 | B | Coop 0.8,-0.9; 0.8,-0.9; -0.5,0.5; -0.5,0.5; 0.8,-0.9; -0.5,0.5; -0.5,0.5 | Both agents execute their own preferred action and feeling and no mutual empathic understanding exists, no solid joint decision has been formed. In comparison with a weaker connection between metaphor and ownership state, for each agent, all self ownership states develop stronger, other ownership states develop less strong due to more influence of the competitive metaphor. In this scenario, the decisions are formed more quickly (even though not joined), and the agents are more competitive, e.g. they execute their own preferred action more strongly and the other action less strongly than with a weaker metaphor-ownership connection. |
7.5. Discussion

In this paper, the influence of cognitive metaphor on joint decision making processes was examined using a social agent model. The social agent model presented in this paper was based on mechanisms for joint decision making known from social neuroscience and on cognitive metaphor theory [13], [20], [21], [34], [43]. These are two concepts that have not been combined in a computational model before. Among the core mechanisms adopted are mirror neurons [3], [5], [6], internal simulation [3], [8], [9-11], and ownership states [3], [12], [28], [29]. In the model, the integration of cognitive metaphor in joint decision making was addressed in general, but illustrated for a cooperative and competitive metaphor state in particular. The metaphor states are activated by the activation of certain sensory representations. In turn the metaphor states affect the agent’s ownership states and thus a decision to perform an action or not. It was shown how a social agent model can be used to simulate and analyse different scenarios for joint decision processes influenced by cognitive metaphor. The outcomes were shown for several possible scenarios. Although the outcomes of the different situations currently simulated already show that metaphors have an influence on the joint decision making process, the model still needs to be tested with more scenarios, for example with different metaphors.

The presented model could have many possible uses. For example, it could be helpful in the design of human-like virtual agents for simulation based training, like a virtual agent helping a human in training (group) decision tasks. Another possible use for a more specific task is a virtual agent that helps partners as a mediator how to learn how to make solid joint decisions, or to help business to come to a most valuable deal with another business.

In an extension of this research several options could be considered. For example, the current model could be adapted to enable the use of multiple metaphors. Furthermore, the way that metaphors are activated in the model could be extended by including other external influences in the activation process. Finally, the influence of disorders on the use of metaphor in decision making would be an interesting concept to examine further.

References

Chapter 7


Part IV:

Application of Integrative Computational Models in Complex Real-World Domains: Aviation and Domestic Energy Management

As part of the process of designing and developing cognitive models it is important to evaluate and explain how to apply them in real world situations. Two domains were selected for this purpose: aviation and energy management. For the aviation domain situation awareness was simulated through a dynamic cognitive model that was developed. These simulations cover a number of different situations: poor perception (due to failure to correctly perceive information), incorrect comprehension (due to failure to rationally comprehend the situation), incorrect projection (due to failure to project a future situation properly), and a conflict between what is predicted and what actually occurs. For the energy domain, some research was conducted to evaluate the use of the type of models developed in heat pump related energy management. Therefore, an analytical model for mathematical analysis of smart daily energy management for air to water heat pumps was developed and it was separately validated with some real data too. Having this analytical model (with separate validations) and a cognitive model for action selection, these two were integrated to provide insights for cognitive driven energy management choices. In addition, in all applications of cognitive models an important challenge is parameter estimation. Therefore, an improved parameter estimation method was developed that is particularly useful for dynamic cognitive models for which not much empirical data can be collected.
Chapter 8

A Parameter Estimation Method for Dynamic Computational Cognitive Models\textsuperscript{1}

Dilhan J. Thilakarathne

Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.

e-mail: d.j.thilakarathne@vu.nl

Abstract: A dynamic computational cognitive model can be used to explore a selected complex cognitive phenomenon by providing some features or patterns over time. More specifically, it can be used to simulate, analyse and explain the behaviour of such a cognitive phenomenon. It generates output data in the form of time series which can only be partially compared to empirical knowledge. This leads to a challenging problem to estimate values of the parameters of the model representing characteristics of a person. A parameter estimation approach for dynamic cognitive models is presented here by combining improved Particle Swarm Optimization (PSO) and Constraint Satisfaction (CS) methods. Having collected the key features of behaviour of a phenomenon, those are translated into a set of constraints with parameters that will be solved through an improved agent based PSO technique. Through this, within PSO each agent explores the complex search space while communicating the quality of a local parameter value vector relative to their current global best solution as a swarm (through cooperation and competition). This is performed in tournaments and results of each tournament are combined to address the premature convergence issue in PSO.

Keywords: Parameter estimation, Particle swarm optimization, Constrain satisfaction, Cognitive modelling.

\textsuperscript{1} This chapter was published as:
http://doi.org/10.1016/j.procs.2015.12.178
8.1. Introduction

With the developments in brain-imaging and recording techniques, more and more detailed information on various brain processes becomes available and this contributes a strong increase in the development in cognitive modelling [1]. It has been found that more than 80% of published articles in theoretical journals in cognitive science are about cognitive modelling [2]. Furthermore, computational modelling is considered to be an important pillar for the development of cognitive science and its related disciplines [3]–[5]. In principle, in cognitive modelling, a phenomenon of human cognition (or behaviour) is represented as (computational) mathematical models to understand the causality with exposed adjustable parameters that can be estimated from empirical data to align with reality [2], [6]. The analogy between the actual human cognition and empirical evidences collecting through brain imaging and recording techniques occurs at two different levels of abstraction. Pragmatically, this nature of observed output (in most cases these will be in the form of features or patterns over time that explain the characteristics in non-quantitative, discrete forms) is very useful to understand the workings of the highly complex human brain. However, this makes it a non-trivial challenge to validate the behaviour of such a model, due to not having an exact quantitative outcome about features or patterns of human cognition. Almost all models developed in science and engineering include parameters (representing characteristics of a person or process) and that is a key driving element to bridge the gap between what is predicted and what is observed [7]. The most generic approach in parameter estimation is systematically changing the parameter values such that the error between predicted and observed comes closer to zero. This strategy is difficult to directly adopt for cognitive models and especially for dynamic cognitive models (see [6]).

Given the nature of available empirical data, first it is essential to translate these (incomplete) output data into a more well defined and quantifiable form. One of the most promising generic representations for this is to express those data as (temporal) constraints. Dynamic features and/or patterns can be easily represented as set of constraints using formal languages. Having a set of such constraints, the parameter estimation process condenses to a problem of Constraint Satisfaction (CS). A large number of problems in computing domains are represented directly or indirectly as Constraint Satisfaction Problems (CSP) [8], [9] and most of these problems belong to the class of NP-complete problems, including most of cognition related problems (see [10], [11]). Therefore, it is a challenge to develop a technique to solve a problem of constraint satisfaction that is capable of finding a solution within a reasonable computational expense. Stochastic and heuristic based searching techniques are commonly used for this purpose and each having specific advantages and disadvantages. Particle swarm optimization (PSO) technique is attracting the
attention of many researchers in many domains due to its efficiency and effectiveness. Main advantages of the PSO methodology are simplicity in calculation, easiness in implementation, comparatively little number of parameters, and free from derivatives [12]. In this paper, a generic approach for parameter estimation in dynamic computational cognitive models is presented by combining CS and improved PSO techniques. It is illustrated for a complex cognitive model addressing action awareness.

8.2. Theoretical basis of CS and PSO

CS and PSO are well known techniques in artificial intelligence (PSO is relatively new) for many problems and both are having well-defined mathematical basis.

8.2.1. Constraint Satisfaction

A constraint satisfaction problem is defined by a finite set of variables \( X = \{ x_1, x_2, \ldots, x_n \} \) (each variable \( x_i \) has its own domain \( D_x \) that includes values \( \{ v_{1}, v_{2}, \ldots, v_{m} \} \)) and a set of constraints \( C = \{ c_1, c_2, \ldots, c_p \} \) (each constraint \( C_i \) involves some subset of variables). The problem is said to be solved if it is possible to find an assignment of values to variables such that it satisfies restrictions imposed by all the constraints [9]. Even though the representation of a CSP is very simple it is a non-trivial task to find a solution tuple depend on the complexity of a given problem. Constraints restrict the values that each variable can simultaneously holds and the challenge is to identify a searching technique to find a global assignment to all the variables that satisfies all the constraints with both efficiently and reasonably less computational expense (for such techniques see [8], [9]).

8.2.2. Particle Swarm Optimization

PSO is a relatively new population based hybrid (both heuristic and stochastic) search/optimization technique that can be used to find (approximate) solution(s) in effective and efficient manner. PSO is first introduced by Kennedy and Eberhart [13], and it emulates a social system that can be observed in a flock of animals that have no leaders. These flocks achieve their goals through communication with others by sharing information about current situation, and therefore group will condense to a position which has a better solution. This phenomenon continues until the best conditions or a situation is discovered. PSO also consists with swarm of particles where each particle holds a position that represents a solution (in a search space) and has a velocity. Particles are moved based on three configurable factors: social, cognitive, and speed [12], [14], [15]. Social factor influences the convergence towards the best solution discovered so far by a particle of the swarm,
and cognitive factor influences each particle on its best position discovered so far, and speed factor delimit the movement. Therefore, always a movement is a resultant of these three factors. These three factors again coupled with another parameter set such that convergence can be biased to social, cognitive, speed or any permutation in these three (see [14]). This behaviour can be represented in two simple equations:

\[ v_i(t + 1) = \omega v_i(t) + c_1 r_1(t) \left( P_i - X_i(t) \right) + c_2 r_2(t) \left( P_g - X_i(t) \right) \]  

\[ X_i(t + 1) = X_i(t) + v_i(t + 1) \]  

Where \( v_i(t) \) represents velocity of the \( i \)th particle, \( P_i \) represents the best previous position of the \( i \)th particle, \( P_g \) represents the best position discovered so far among all the particles in the swarm, \( X_i(t) \) represents the position of the \( i \)th particle. \( X_i(t) = (x_{i1}, x_{i2}, ..., x_{iD}) \) (where \( D \) is the dimension of the search space), \( c_1 \) and \( c_2 \) are positive constants called acceleration coefficients (\( c_1 \): cognitive factor of the particles, \( c_2 \): Social factor of the particles), \( r_1 \) and \( r_2 \) are two random numbers in the range [0,1], and \( \omega \) is the weight of inertia for the velocity.

In each iteration, each particle evaluates the quality of its current position. This evaluation can be combined with the CSP. In PSO, a position of a particle is a (probable) solution for the finite set of variables \( X \) mentioned in the CSP. Each particle shares its information with the swarm and updates the new position as in the equations (1) and (2). PSO has a major limitation due its premature convergence problem [15], [16]. As there is a biasness towards to the global best solution, having discovered a local optimum all agents will scatted towards to that. This can be handled by minimizing the influence towards the global best solution by providing a very small \( c_1 \) value but a large \( c_2 \) value (see [15]). Then the particles are having a problem of getting converge and they just explore around their local best position than exploring the search space proactively. Therefore, this premature convergence is not easy to handle and various approaches were presented in the literature (see [16]). For this paper, an improved PSO algorithm is scrutinized by considering the concept of tournament based selection. In each tournament, particles are trying to find a solution that satisfies all the constraints but in most cases, they may trap in a local optimum. If it is identified that the swarm is taped in a local optimum (with the same global best solution for ‘m’ number of consecutive iterations) the global best solution of that tournament stores and initiates a new tournament. After ‘n’ number of consecutive tournaments, swarm has identified n best local optimums (in the meanwhile, there is a possibility of finding the best solution and in such a case process is terminated). Having n number of local optimums, each particle is replaced with a local optimum and let swarm to explore the search space. This process continues until finding a global solution that satisfies all the constraints.

Pseudocode of this improved PSO is provided in Pseudocode 1. There are some rule of thumbs to select parameter values for the PSO algorithm [12], [15]. Number of
particles \( n \) to be selected is normally in the range of \([20, 50]\), \( \omega \) is advised to be in the range of \([0.4, 1]\), and the total of \( c_1 + c_2 \approx 4 \).

### 8.3. A Computational Cognitive Model of Action Awareness

Action awareness is a complex cognitive phenomenon that covers the questions how action selection and execution contribute to the awareness and/or vice versa. Some evidence lead to a hypothesis that awareness of action selection is not directly causing the action execution (or behaviour) but comes afterward as an effect of unconscious processes of action preparation \([17]–[21]\). In contrast, another hypothesis claims that both predictive and inferential processes related to the action preparation and execution may contribute to the conscious awareness of the action and furthermore, this awareness of an action is a dynamic combination of both prior awareness (through predictive motor control processes) and retrospective awareness (through inferential sense-making processes) relative to the action execution \([22]–[23]\). By integrating these evidences with other necessary processes that contributes

---

**Pseudocode 1**: Pseudocode of this improved PSO

1. Initialize \( \omega, c_1, c_2, n, m \)
2. For \( i=1 \) to \( n \) do
3. Initialize each \( X_i \) with random values for variables (bounded by given max and min);
4. Initialize each \( v_i \) randomly;
5. End for
6. Initialize an array of tournaments’ solutions (tSolutions) with size \( n \)
7. Do
8. For each particle
9. Calculate fitness value
10. If the fitness value is better than its personal best \( (P_i) \)
11. set current value as the new personal best
12. End
13. Choose the particle with the best fitness value in the swarm as current global best
14. If the current global best value is better than its global best \( (P_g) \)
15. set current value as the new global best
16. Else
17. If convergence stays with the same \( P_g \) value for ‘m’ consecutive rounds
18. add the \( P_g \) of current tournament into tSolutions
19. If the no. of local solutions in lSolutions is \( n \)
20. replace all current particle from particles in lSolutions and GOTO 8
21. Else
22. update each particle with random \( X_i \) and \( v_i \) values and GOTO 8
23. End
24. End
25. End
26. For each particle
27. Calculate particle velocity according to equation (1)
28. Update particle position according to equation (2)
29. End
30. While maximum iterations \( (n) \) or minimum error criteria is not attained
for awareness of action, a computational cognitive model was developed that published in [24]. The same model is used in this paper for the proposed approach. Detailed information of cognitive, affective and behavioural neuroscience literature and a detail model description are not presented in this section but only an overview (for more details see [24]).

Fig. 1 presents the cognitive model developed for action awareness and Table 1 presents the abbreviations used in that. Model uses three world states (WS) as inputs: stimulus $s_{ki}$, context $c_{ki}$, and effect $b_{ki}$. In addition, it includes two outputs to interact with the environment: $EA(a_i)$ and $EO(a_i, b_i, s_{ki}, c_{ki})$. The input world states WS($s_{ki}$), and WS($c_{ki}$) lead to sensor states SS($s_{ki}$), and SS($c_{ki}$), and subsequently to sensory representation states SR($s_{ki}$), and SR($c_{ki}$), respectively. A sensory representation may trigger one or many action preparations: PA($a_i$), in parallel and in unconscious (or automatic) form. The brain will evaluate the effect of each relevant action preparation by comparing the feelings: $F(b_i)$, associated to each individual valuated effects (without actually executing them through the body loop) as proposed by Damasio [25], which is presented as “as-if body loop” ($PA(a_i) \rightarrow SR(b_i) \rightarrow F(b_i)$) in Fig. 1. The simulated option that has the strongest valuated feeling performs as a GO signal through the body loop and else are NO-GO options. Each $PA(a_i)$ state suppresses its complementary options $PA(a_j)$ for $j \neq i$ (as shown in dotted looped red arrow in Fig. 1) proportional to the accumulated strength of that option. This behaviour is in line with the explanation for the lateral inhibition in [26] and contributes to further strengthen the action selection process. Therefore, naturally the

![Figure 1](image_url)
strongest internally satisfied option (which is exceeding a threshold value) will become selected as a result of the unconscious action selection process.

This model further includes action ownership states. Action ownership is a useful concept, which is mainly important to differentiate in how far a person attributes an action to him or herself, or to another person (see [27]). Model provides qualitative analysis on cognition before and after action execution (together with prior and retrospective awareness states). The ownership state includes information of inputs, action preparation, and predicted feeling of the preparing action option. This integrated information is used as the gateway to develop action awareness. Similar to ownership, awareness also separated into prior awareness and retrospective awareness as suggested by Haggard and co-workers [22], [23]. For each ownership state an associated awareness state may (or may not) emerge. Awareness states play a higher order cognitive role. The direct links from ownership and feeling to awareness state realise bottom-up activation (see [28], [29]). Conversely, the effects of awareness states on other states realise top-down activation, which is considered to be a conscious or intended process (see [30]–[32]). Therefore in the presented model PAwr(ai, bi, ci, sk) is only affected by PO(ai, bi, ci, sk) and F(bi). This is useful to model the idea of Benjamin Libet: brains initiate voluntary movements before we are aware of having decided to move [17]–[21]. Moreover PAwr(ai, bi, ci, sk) affects PA(ai) and EA(ai) and this is reflects the idea of Haggard and co-workers: there may be an impact from this subjective awareness state on action execution [22]–[23]. By this PAwr(ai, bi, ci, sk) to PA(ai) link the agent can inject some bias to the current unconscious process through awareness. This may strengthen a weaker action option and improve the predictive feeling of that option (which may lead to getting it executed) [33]. Furthermore, in this model PAwr(ai, bi, ci, sk) can also directly strengthen the action execution state. By considering the proposed model as an agent, a particular behaviour trace is expected which is considered as main characteristics

| Table 1: Nomenclature for Fig. 1 |
|-----------------------------|---------------------------------|
| WS(W) | world state W (W can be either: context c_k, stimulus s_k, or effect b_i) |
| SS(W) | sensor state for W |
| SR(W) | sensory representation of W |
| PA(ai) | preparation for action a_i |
| F(bi) | feeling for action a_i after as-if loop or action execution |
| EA(ai) | execution of action a_i |
| PO(ai,bi,ck,sk) | prior ownership state for action a_i with b_i, c_k, and s_k |
| RO(ai,bi,ck,sk) | retrospective ownership state for a_i with b_i, c_k, and s_k |
| PAwr(ai,bi,ci,sk) | prior-awareness state for action a_i with b_i, c_k, and s_k |
| RAwr(ai,bi,ci,sk) | retrospective-awareness state for action a_i with b_i, c_k, and s_k |
| EO(ai,bi,ci,sk) | communication of ownership and awareness of a_i with b_i, c_k, and s_k |
of the simulation example for parameter estimation. Expected trace is shared in an external appendix\(^2\) (from [24]). The trace provides information to extract some important patterns to isolate the behaviour. The order of activation of the states is a very important feature. Furthermore, it is expected to be having a dip in the sensory representation and feeling in-between predictive and inferential representations [34]–[36]. In addition, the maximum strength of each state should exceed some threshold value to highlight the contribution of each state as the literature presented in [24].

The proposed model is compiled to a computational model from a dynamic system perspective. The detail information of the model compilation and its parameter information (i.e., steepnesses: \(\sigma\), thresholds: \(\tau\), weights: \(\omega_k\), update speed factors: \(\gamma\), and step size: \(\Delta t\)) can be found in [24]. In our previous work for this, a systematic analytical driven parameter estimation method was used with realistic results (see [24]). That work includes 8 scenarios and only the first scenario will be considered for the scope of this paper. In [24] first scenario is focused on the interplay between conscious and unconscious processes on action selection where the prepared action has satisfactory predicted effects and therefore is executed; in this case both prior and retrospective awareness states occur. The behavioural results for first scenario in [24] is not presented in here. This result is used to compare the quality of proposed parameter estimation method through CS and PSO.

### 8.4. Parameter Estimation with PSO

The model presented in Section 3 together with computational basis is considered for PSO based parameter estimation approach. The system includes 17 steepnesses, 17 thresholds, 39 weights, 2 update speed factors, and 1 step size. Each of these considered as variables and assigned domains as required in the CSP. Nevertheless, heuristic knowledge is used for assigning domain values for each variable to make the complexity less intractable and to prune the search space to increase the performance. A pre analytical analysis conducted to identify variables where its value can be pre determined. For example, WS(\(s_k\)), and WS(\(c_k\)) are the input states of the model and obviously \(\sigma\) values should be almost zero and therefore value 0.01 is used for these together with value 1.0 for \(\tau\). Furthermore, state F(\(b_i\)) affects only from SR(\(b_i\)) and therefore, it is obvious that from causality perspective F(\(b_i\)) should activate as soon as SR(\(b_i\)) has started to emerge. Therefore, \(\sigma\) value of SR(\(b_i\)) should also be almost zero and value 0.01 is used for this. In additionally, as SS(\(s_k\)), SS(\(c_k\)), SR(\(s_k\)), and SR(\(c_k\)) states includes direct representation of the inputs through weights \(\omega_5\), \(\omega_6\), \(\omega_8\), and \(\omega_9\) (see [24]); these assigned with value 1.0 (this is an implementation decision, if only a portion from input strength should be used for this then it can be < 1.0). Furthermore, according

\(^2\) [http://www.few.vu.nl/~dte220/BICA2015_Appendix.pdf](http://www.few.vu.nl/~dte220/BICA2015_Appendix.pdf)
to the behaviour necessary for the highlighted scenario in the previous section, it is clear that weight values attached to some states should have very high values or above average values or very small values or negative values (high, average or small). Therefore, than using a domain range for weights 0 to 1 or 0 to -1, different ranges are used for small, average, and large (e.g., -0.7 to -0.4, -1.0/-0.9 to -0.6, 0.1 to 0.5, 0.4 to 0.7, 0.6 to 0.9/1.0). For steepness and threshold values also rather than using a generic range, custom ranges used with prior experience. From analytical perspective, in a model like this the states which are activates at the beginning normally shouldn’t have very strong steepness and threshold values as they are directly coupled to the inputs and having strong weight values naturally more strength will be collected. Nevertheless states come later (especially after action execution) it is important to have very strong values for steepness and threshold variables. All value ranges used for variables available in an external file3 (this includes the initial velocity values used in the PSO algorithm too). In addition to the variables and there domain values, a set of constraints (C1 to C5) also defined based on the neurocognitive evidences.

C1. State activation order should perceived: WS(sk) & WS(ck) → SS(sk) & SS(ck) → PA(ai) → SR(bi) → F(bi) → PO(ai, bi, ck, sk) → PAwr(ai, bi, ck, sk) → EA(ai) → WS(bi) → SS(bi) → RO(ai, bi, ck, sk) → RAwr(ai, bi, ck, sk) → EO(ai, bi, ck, sk)

C2. Each state value should converge to zero after retrospective effects
C3. All states except SR(bi) and F(bi) should only provide one peak behaviour
C4. States SR(bi) and F(bi) should provide two peak behaviour where the value at first peaks should be considerably small respective to the value at the second peak
C5. Peak value (max) of each state should be equal or exceed a given value: 1, 1, 0.55, 0.55, 0.50, 0.50, 0.65, 0.50, 0.65, 0.50, 0.65, 0.65, 0.50, 0.50, 0.65, 0.65, 0.70 (hear the order is the same order of states in C1)

These constraints are noted as the most influential restrictions to have necessary generic behaviour. There is a trade off between minimum and maximum constraints required for a model to isolate a generic parameter value set to mimic its behaviour with reality (see [8]). Each constraint associated with a negative error value such that if all constraints are satisfied the value of the error is zero. For the constraint C1, -13 is assigned as the total error and if two consecutive positions are as in the order then +1 will be assigned (therefore, if all the states are in the expected order the error is 0). For the constraint C2, -17 is assigned as the total error and if a selected state is converging to an activation strength of 0 after the retrospective effect then it will get +1. For C3 and C5 also the same above rule applied for each

3 http://www.few.vu.nl/~dte220/IAT15_ModelVariablesS1.xml
state. For the constraint C4, -2 is assigned as the total error and if a selected state is showing a two peak behaviour it gets +1. Having represented the proposed model as a CSP, it is solved by using the proposed PSO algorithm. For this purpose, following configuration parameters are used for the PSO equations (1 & 2) presented in Section 2: number of particles (n) is 30, weight of inertia for the velocity (ω) is 1, local acceleration coefficient (c1) is 2, global acceleration coefficient (c2) is 2. These selected parameter values are according to the guidelines given in [12], [15]. Having these parameters for the PSO equations (1 & 2) the model is implemented in the Java language and passed the initial values in a XML data file. The pseudocode presented in the Section 2.2 is used for the PSO implementation. First 30 particles were randomly initialized for the 73 variables (17 steepnesses, 17 thresholds, 39 weights) under the domain ranges provided for each variable. For the low and fast update speed factors (γ) values 0.6 and 0.7 are used respectively together with value 0.25 as the step size (Δt). Having the same value for local acceleration coefficient (c1) and global acceleration coefficient (c2) the swarm is not biased to local or global exploration but due to the random values assigned for r1 and r2 (see equations 1 & 2) in each iteration swarm may randomly select a local exploration or global exploration.

Having 30 random particles in the search space, they evaluate the quality of current positions individually by mapping with the five constraints. If the current local best position is better than the past best local position visited (only after the first iteration) it will be replaced by the current. Having identified the quality of current positions by each particle, they share the information with each other to find the current global best position. If the current global best position is better than the previous global best position (only after the first iteration) as a swarm, then all particles change their global best to the current global best. In additionally, each particle changes its current position towards to the resultant of global best position (i.e., social aspect), local best solution (i.e., cognitive aspect), and speed factor. This process executes as a tournament. In each tournament, if any particle’s position able to satisfy all the constraints, then the process will be terminated and the variable values of that particle’s position will be used as the parameter values in the model. If a tournament fails to find a particle that satisfies all the constraints then the global best position of the swarm at the end of that tournament will be separately saved and new tournament will be initiated with random positions for all particles. It is decided that a tournament is failed if the same error value for constraints holds for 15 iterations continuously. After a 30 tournaments if the system is unable to find a solution the captured best global positions in each tournament will be used and create a swarm of 30 particles and let the system to converge. In this case, there is a very high probability to have many particles with the same low error values (most probably very closer to the 0) and therefore, speed of particle movements will be
limited but more local explorations performs. This particular improvement introduced as a solution for premature convergence problem. A major drawback in PSO is its premature convergence ([15], [16]) and putting quality particles after 30 tournaments improve the convergence speed a lot and due to this elicit selection feature, the system will not discard better solutions for future explorations. This iterative process continues until a particle is obtained a position that satisfies all the constraints. Fig 2(a, b, & c) shows the behaviour patterns obtained from the proposed PSO algorithm for the mentioned cognitive model. All the results are completely satisfied with all the constraints and align with the evaluation in [24]. There are three different solutions included in the Fig 2(a, b, & c), which are provided by three different executions. Nevertheless, all the results are fully satisfied by the expectations (constraints).

8.5. Discussion

Parameter estimation is a challenging task in many application domains, including the dynamic cognitive modelling domain. Nevertheless, it is difficult to adapt common parameter estimation algorithms for cognitive models where usually there are only limited empirical data available (a.o. due to limitations in neuro-imaging techniques and the complexity of the human brain). Moreover, most of the behaviours are presented in the form of features or patterns over time that are explained in non-quantitative and discrete forms. This paper presents a solution for such cognitive models using an improved PSO algorithm. The complex cognitive

Fig. 2(a): Cognitive behaviour obtained through PSO based parameter estimation algorithm for executing an action with ownership and awareness
model used in this paper has been published in [24] together with eight scenarios to validate the model behaviour. Only the first simulation of that work was considered in this paper as an illustration for parameter estimation, and acceptable results were achieved. According to these results it is confirmed that this new approach is suitable for parameter estimation for these types of complex cognitive models. Initially for this purpose a standard parameter estimation approach was used without introducing a tournament based improvement. It was found that then the method is unable to converge to a solution and get trapped with local minima all the time. When exploring approaches to eliminate premature convergence issues, there are some techniques available (see [16]). Nevertheless, all of these techniques are having limitations and there are some technical problems to adapt those techniques to the cognitive domain.

The Multi-Swarm PSO (MSPSO) technique ([37], [38]) was chosen as a suitable solution for this issue and it has many supporting features for parameter estimation in dynamic cognitive models. The most common approach in the MSPSO is to partition the current swarm into several sub-swarms. Nevertheless, due to computational problems in parallel searching and difficulties in forming sub-swarms as equally representing the partitions of the search space, this is also not providing promising results. Therefore, in this paper a new approach is presented by introducing a tournaments based approach. Therefore, no parallel searching is required and having a strong random particle generator it will guarantee a quality distribution of particles in the search space. It is a future work to conduct a systematic comparison this with MSPSO. According to the results in Fig 2(a, b, & c) it is clear that this approach is not just giving one solution but multiple near optimal solutions. This seems like a
problem as it is not able to properly represent a generic behaviour of the cognitive model. This problem is addressed in our previous work [24] by introducing 8 different scenarios (which are interrelated from a functional point of view) and finding a unique parameter value set that captures the behaviour of all the simulations. Having interrelated scenarios, the goal is to find a global parameter value set that satisfies all the of those individually (for this it is necessary to have many near optimal solutions at the beginning, but later the system will converge to a one solution due to the feature of interrelatedness among scenarios). This is a future work for this technique to combining the procedure mentioned in [24] for this PSO algorithm. In additionally more validations are required for this approach and it is essential to do more thorough analysis on effects of changing number of particles, changing the parameter values of the PSO algorithm, removing/diluting heuristic knowledge used to set domain ranges, and measuring average running time. Also, having thoughts to present this cognitive model as a generic framework to use as an experimental workbench on action selection related cognitive phenomenon, it is important to have a feature that the modeler does not need to do any programming to obtain the behaviour. The current implementation only needs an input in a XML file and results will be generated except representing the constraints in Java. Having a formal representation to present the constraints will eliminate this limitation and a predicate logic based representation will be used for this in future work.

Fig. 2(c): Cognitive behaviour obtained through PSO based parameter estimation algorithm for executing an action with ownership and awareness
Acknowledgment

I wish to thank Prof. Jan Treur at VU University Amsterdam, for his great support and supervision in all the phases of this work.

References

Chapter 9

Modelling of Situation Awareness with Perception, Attention, and Prior and Retrospective Awareness

Dilhan J. Thilakarathne

Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.
e-mail: d.j.thilakarathne@vu.nl

Abstract: Human awareness under different circumstances is complex and non-trivial to understand. Nevertheless, due to the importance of awareness for safety and efficiency in many domains (e.g., the aviation domain), it is necessary to study the processes behind situation awareness, to eliminate possible errors in action selection that may lead to disasters. Interestingly models for situation awareness have been presented, mainly from an ecological psychology perspective, but they are debatable with respect to the latest neurocognitive evidences. With the developments in brain imaging and recording techniques, more and more detailed information on complex cognitive processes becomes available. This provides room to further investigate the mechanisms behind many cognitive phenomena, including situation awareness. This paper presents a computational cognitive agent model for situation awareness from the perspective of action selection, which is inspired by neurocognitive evidences. The model integrates bottom-up and top-down cognitive processes, related to various cognitive states: perception, desires, attention, intention, (prior and retrospective) awareness, ownership, feeling, and communication. Based on the model, various cognitive effects can be explained, such as perceptual load, predictive processes, inferential processes, cognitive controlling, unconscious bias, and conscious bias. A model like this will be useful in domains that benefit from complex simulations of socio-technical systems (e.g., the aviation domain) based on computational models of human behaviour. In such domains, existing agent-based simulations are limited, since most of the agent models do not include realistic nature-inspired processes. The validity of the model is illustrated based on simulations for the aviation domain, focusing on a particular situation where an

1 This chapter was published as:
http://doi.org/10.1016/j.bica.2015.04.010

This Journal paper is an extended work of following two conference papers:
http://dx.doi.org/10.1007/978-3-319-19066-2_9

http://doi.org/10.1007/978-3-319-09891-3_42
agent has biased perception, poor comprehension, habitual driven projection, and conflict between prior and retrospective effects on action execution.

**Keywords**: Situation Awareness, Prior and Retrospective Awareness, Bottom-Up, Top-Down, Cognitive Modelling.

### 9.1. Introduction

The relation between human awareness and action selection is a complex issue, which is the subject of much debate and provides many challenges for further research. Nevertheless; due to the developments in brain imaging and recording techniques, the insight in human brain processes is growing rapidly, which contributes to an improved quality of relevant data and to the development of new methods to explore this most complex system within human anatomy through different dimensions. Human cognitive processes are often grouped into conscious (i.e. accompanied with awareness) and unconscious processes. The understanding of the interplay between conscious and unconscious processes associated with action selection and related phenomena has much improved, especially thanks to the experimental framework proposed by Libet, Gleason, Wright, & Pearl (1983) and later improvements made to it. In the literature, bottom-up cognitive processes have been mapped to unconscious action formation, whereas top-down processes have been related to the conscious action formation (cf. Moore & Haggard, 2008; Engel, Fries, & Singer, 2001; Haggard, 2008; Kiefer, 2007); it seems our action selection process initiates from unconscious phenomena, and that later we develop the conscious experience of this action selection. The unconscious neural activations in the brain seem to be a result of habitual tasks, through the effects of prior learning, which can be automatically activated when a relevant stimulus is perceived (Monsell, 2003). Nevertheless, conscious awareness of action selection also plays an important role and the influence of predictive and inferential processes of action execution has been highlighted by Haggard and co-workers, providing a working mechanism for this process (cf. Moore & Haggard, 2008).

Situation Awareness (SA) can be considered as a subjective quality or interpretation of the awareness of a situation a person is engaged in. When a person is engaged in a situation based on the information that he/she perceives, the attention that is allocated to that information based on his/her subjective desires will develop his/her subjective awareness of the situation. This is the reason why different individuals may have different interpretations of the same situation. The
correctness of SA is always relative and its quality can be analysed when a task is performed with an expert critiquing as a benchmark. Due to this complexity and subjective nature of SA, the concept has received many definitions in the literature and according to (Dominguez, 1994) there are more than fifteen definitions about SA; among those, the definition proposed by Endsley (1988) became the most widely used. According to Endsley, SA is:

"the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1995, p. 36)

Based on this definition, Endsley highlighted three elements as the necessary conditions for SA; these are three levels of which one is followed by the other, in order to develop complete (subjective) awareness. These three elements are the following:

1. Level 1: perception
2. Level 2: comprehension and
3. Level 3: projection

Furthermore, it has been found that, based on safety reports in the aviation domain, 76% of the errors related to SA were because of Level 1 (i.e., failure to correctly perceive information), 20.3% were Level 2 errors (i.e., failure to comprehend the situation), and 3.4% were Level 3 errors (i.e., failure to project situation into the future) (Endsley & Garland, 2000; Endsley, 1999). Hence, this statistical information provides an indication of the relative importance of these three aspects of SA. It has been noted that human error is a significant factor (71%) for accidents in the aviation domain, and among those, 88% of the accidents are directly related to the SA problems (Shuang, Xiaoru, & Damin, 2014). These statistics highlight the importance of understanding SA from a more cognitive perspective.

Furthermore, Endsley has indicated how attention, goals, expectations, mental models, long-term memory, working memory and automaticity contribute to situation assessment in terms of cognitive processes (Endsley, 1995; Endsley, 2000). The following summary from Endsley provides some useful indications of how this definition (through her model) can be related to the neurocognitive literature (presented in Section 2):

“To summarize the key features of SA in this model, a person’s SA is restricted by limited attention and working memory capacity. Where they have been developed, long-term memory stores, most likely in the form of schemata and mental models, can largely circumvent these limits by providing for the integration and comprehension of information and the projection of future events (the higher levels of SA), even on the basis of incomplete information and under uncertainty. The use of these models depends
on pattern matching between critical cues in the environment and elements in the model. Schemata of prototypical situations may also be associated with scripts to produce single-step retrieval of actions from memory. SA is largely affected by a person’s goals and expectations which will influence how attention is directed, how information is perceived, and how it is interpreted. This top-down processing will operate in tandem with bottom-up processing in which salient cues will activate appropriate goals and models. In addition, automaticity may be useful in overcoming attention limits; however, it may leave the individual susceptible to missing novel stimuli that can negatively affect SA” (Endsley, 1995, p. 49)

Though there is a positive analogy between Endsley’s model and its psychological basis, from a dynamic perspective the model leaves room for questions. In particular, she highlights that a person will first develop situation awareness, only then decision making will follow, and that finally the selected actions will be performed:

Environment → Situation Awareness < Perception → Interpretation → Projection > → Decision Making → Performance

Nevertheless, this particular linear transformation is not supported by the latest viewpoints in neuroscience; instead, situation awareness, decision making, and performance of selected actions are viewed as one compound process in which all sub-processes dynamically interact, striving for actions with an optimal result. Endsley has explicitly separated situation awareness from the process itself, which she calls situation assessment (Endsley, 2000) (in her terms, SA is a product and situation assessment is the process to make decisions which follow up on the developed SA). Having this fundamental concern, there is a necessity of better explaining SA in terms of current evidence from a cognitive science perspective, in such a way that the explanations can be used in relevant application domains (such as the aviation domain), for instance by simulating complex situations in a more realistic and detailed manner.

In addition to Endsley’s model, there are a few more influential works on SA. Among those, the perceptual cycle (or the perception/action cycle) proposed by Neisser (1976) was adopted by many researchers (Adams, Tenney, & Pew, 1995; Durso & Dattel, 2004; Klein, Phillips, Rall, & Peluso, 2007) to model SA. The perceptual cycle includes three elements: object (i.e., current information in the environment), schema (i.e., knowledge developed through training or exposure, such that the operator knows what to do), and exploration (i.e., the operator’s ability to search in the environment after an action). This is a cyclic process:

- An object (or stimulus) triggers attention to bind relevant schema element(s)
- The activated schema motivates the operator (or agent) to perform an action in the environment
Due to the changes made in the environment, the operator explores the environment such that new objects or stimuli are discovered.

This triggers a necessary attention shift, to select a related schema.

The cycle continues ...

This model relates to Endsley’s model in many ways, but the main distinction is that SA is described not only as a product but both as a product and a process. Therefore, when modelling SA, it is important to also consider aspects like action prediction and the observation of the effects of executed actions. The ideas proposed by Haggard and co-workers on action awareness (Haggard, 2008; Haggard & Clark, 2003) are in line with this idea. They have proposed that awareness of an action is a dynamic combination of both prior awareness (i.e., awareness of action prediction) and retrospective awareness (i.e., awareness of the effects of an action), through predictive motor control and inferential sense-making relative to the action execution, respectively. Therefore, it is possible to integrate other theories of SA with more recent neurocognitive evidences, together with Endsley’s taxonomy to uplift the current state of the art of SA models.

Hence, this paper presents a computational agent model for SA, with an emphasis on biased perception, bottom-up and top-down processes, cognitive controlling, prior and retrospective processes. The model is based on evidences from cognitive, affective, and behavioural sciences, combined with insights derived from Endsley’s and Neisser’s models. This paper is an extension of previous works (Thilakarathne, 2014a; Thilakarathne, to appear), with necessary adaptations to include prior and retrospective processes. In particular, a computational model for SA was presented in (Thilakarathne, 2014a), driven by the interplay between bottom-up and top-down processes together with processes and states such as: perception, attention, intention, desires, feeling, action preparation, ownership, and communication. The main addition in (Thilakarathne, to appear) compared to (Thilakarathne, 2014a) is the inclusion of a phenomenon called perceptual load, which is considered to be the main reason for most of the errors in SA according to Endsley’s taxonomy (Endsley & Garland, 2000). Furthermore, this paper improves the simulation results presented in the previous works and also provides a new simulation that explains the cognitive conflict when there is a mismatch between what is prepared by expecting a particular effect (say $b_1$) and its actual effect after action execution (say $b_2$). The paper is organized as follows: it starts in Section 2 with a summary of relevant evidence from cognitive, affective, and behavioural sciences related to the cognitive states and processes associated with action selection related SA. In Section 3, the proposed model will be explained together with its compilation to a computational model with the necessary mathematical basis. Section 4 presents a validation of the model, particularly providing for simulations. Finally, Section 5 presents a discussion and a future direction for this work.
9.2. Situation Awareness Related Processes Viewed Neurologically, Psychologically and Behaviourally

Cognitive processes of action formation are complicated and on the exact mechanisms involved have not been fully unravelled yet. Nevertheless, due to the developments in brain imaging and recording techniques, researchers are gaining more and more insight in human brain processes. According to those insights, human action selection is for a main part determined by automatic, unconscious processes such as habitual tasks (note that a task will become a habitual task through a learning process, depending on its frequency and recency (cf. Monsell, 2003)). Nevertheless, stimuli received from the environment contain far more information than a person can process in a given time. To cope with this, it seems that two main processes of cognitive controlling play a role, namely bottom-up and top-down control. Processes related to bottom-up effects are more automatic and are mainly data driven, triggered by factors external to the agent (“Humans are agents. That is, they have the capacity to bring about change in the external world through their own goal directed behaviour.” (Moore & Obhi, 2012, p. 547)) such as salient features of a stimulus (Katsuki & Constantinidis, 2014). Top-down effects are more internally guided based on prior knowledge, intentions, and long-term desires and they are internal to the observer and unrelated to the salient features of a stimulus (Kiefer, 2007; Katsuki & Constantinidis, 2014; Awh, Belopolsky, & Theeuwes, 2012; Baluch & Itti, 2011).

When investigating the brain circuits related to human cognition, it seems that these consist of complex loops, rather than linear chains (Haggard, 2008); therefore, higher order coupling among processes has been observed, rather than those processes being categorically independent. Similarly, the bottom-up and top-down processes are also not isolated; instead, there is overlap among these two, even at the neural level and many pieces of evidence have been found that demonstrate the interplay between these two in the context of attention and perception, together with other supportive cognitive states (Kiefer, 2007; Katsuki & Constantinidis, 2014; Baluch & Itti, 2011; Miller & Cohen, 2001; Rigoni, Brass, Roger, Vidal, & Sartori, 2013). Therefore, when an agent perceives salient features of a stimulus unconsciously (through bottom-up processes), that may not always lead to the execution of a habitual task, but may form a more complex cognition as an innate ability by integrating prior knowledge, intentions, and long-term desires together with emerging consciousness (through top-down processes) to make rational decisions.

The bottom-up processes have many relations to perception and emotions, and more details of the cognitive basis of these have been separately presented in previous work (see. Thilakarathne & Treur, 2014). The amygdala was noted as a
key element in bottom-up processes, which include monitoring the salient features in stimuli and projecting them onto higher levels of cognitive processing (the amygdala has connectivity with eight of the cortical areas (Pessoa, 2010a)). Furthermore, it has been observed that the amygdala directly shapes perception when perceiving an emotionally salient stimulus (Pessoa, 2010b). Attention is another important cognitive state that relates to the interplay among bottom-up and top-down processes: in particular through bottom-up attention and top-down attention. In (Katsuki & Constantinidis, 2014) it has been pointed out that the posterior parietal cortex (PPC) and prefrontal cortex (PFC) could be segregated for distinct roles in bottom-up and top-down attentional systems, and the close interaction of these regions with each other is highlighted to explain the constant influence of these two processes to orient the attention necessary for more sophisticated cognitive control processes (cf. Baluch & Itti, 2011; Poljac, Poljac, & Yeung, 2012; Bor & Seth, 2012).

In addition, the prefrontal cortex has long been assumed to play an important role in top-down driven cognitive control, as a temporal integrator. The higher order interconnectivity of the PFC with other cortical and subcortical areas has been interpreted as indicating a process that generates and maintains information when sensory inputs are weak, ambiguous, rapidly changing, novel and/or multiple options exist (Miller & Cohen, 2001; Miller, 2000). Furthermore, neurocognitive evidence for some of the main factors of top-down processes (i.e., intention, attention, subjective desires and awareness) has been presented separately in previous work (Thilakarathne, 2014b). In this context, the anterior insular cortex (AIC) appears to be involved in conscious awareness, accompanied with feelings; and through this humans may gain the capacity to be aware of themselves, others and the environment ((Bud) Craig, 2009). There is a strong correlation between awareness and the intention to act (Haggard & Clark, 2003). Intentional actions are mainly goal driven and therefore, an experience of intentional action includes an implicit feeling that action occurred as a cause of intention to act through the awareness. Furthermore, the activity of the prefrontal parietal network (PPN) has been observed in almost all demanding or novel tasks which are usually attributed to cognitive control (Bor & Seth, 2012). An overlap between attention and conscious awareness also has been suggested, due to the PPN which is evoked for the cognitive functions attention, executive functions, and working memory, and also for conscious awareness (this may highlight the interplay among these and possible influences from conscious awareness to attention; though attention might not be necessary and sufficient in order for conscious awareness to emerge; (cf. Bor & Seth, 2012; Tallon-Baudry, 2012). Furthermore, functional evidences of top-down mechanisms in attentional selection relating to cognitive control have shown
its relation to long-term memory and more specifically to the declarative and procedural knowledge in it (cf. Engel, Fries, & Singer, 2001).

It is important to understand the reasons for poor or incorrect perception. One suitable analogy for this is the literature on distraction, which explains why and how people get distracted from their current task (under a poor ‘Level 1 SA’, agents are unable to switch to proper perception due to the focus developed on selected items or data in the stimulus). Once a person is focusing on something, there are many reasons why he/she may sometimes be distracted and sometimes not. This is an interesting scenario to study the question what are the causes that prevent a person from processing other important cues in the environment (as an example from (Lavie, 2010): people attending to a ball game have failed to notice a woman walking across the pitch and holding up an umbrella). To explain this phenomenon, the load theory of attention and cognitive control in (Lavie & Tsal, 1994) provides detailed information about early versus late selection schemes. Early selection is a perceptual selection mechanism associated with automatic (or passive) behaviour (Lavie, Hirst, de Fockert, & Viding, 2004). The main reason behind this mechanism is the limited processing capacity of our perception under a high level of perceptual load. Because of this, an agent under high perceptual load is unable to shift his or her selection to other salient features in the environment. Instead, when the perceptual load is lower, the agent is capable of perceiving information in parallel (cf. Lavie et al., 2004; Lavie, 2006, 2005). Nilli Lavie in (Lavie, 2010) mentions the following about this:

“The load theory resolves the early- and late-selection debate by combining within one hybrid model the early-selection assumption that perception has limited capacity and the late-selection assumption that perception is an automatic process (in the sense that it is involuntary and so cannot be shut down at will). It follows, then, that tasks involving high perceptual load that engage full capacity will simply leave no capacity for irrelevant distractor perception (leading to an early-selection result). In contrast, in tasks of low perceptual load, spare capacity remaining beyond the task-relevant processing spills over involuntarily to irrelevant distractor processing (leading to late-selection results). The efficiency of late selection (that is, the extent to which distractors that have been perceived can be prevented from gaining control over behavior) depends on the level of load on cognitive-control functions such as working memory. High working memory load during task performance results in greater distractor interference.” (Lavie, 2010, pp. 143–144)

Indeed, there is empirical evidence that shows a competition effect under low perceptual load, but not with high perceptual load (Lavie, 2005). Therefore, when receiving new stimuli with no perceptual load, all features may be processed completely, which will lead to the development of perception on the salient features, while at the same time the perceptual load will increase. Once the perceptual load is high, the agent seems to be unable to pay attention to additional
features. Moreover, perception is related to early selection, whereas cognitive controlling through attention is related to late selection.

It is a long lasting debate among researchers whether or not conscious awareness contributes to the preparation and execution of actions. There are many actions that we perform as habitual tasks, more from the unconscious side of the spectrum. From a SA perspective, it is important to be aware of what one is going to do and how to control this action selection, especially when the task at hand is very important and complex (e.g. landing an airplane). There are many fMRI based studies that show that the brain predicts the outcome of a decision even before the decision reaches awareness (D’Ostilio & Garraux, 2012; Haynes, 2011; Baumeister, Masicampo, & Vohs, 2011; Wegner, 2002; Libet et al., 1983). It has been found that for certain types of actions the decision to perform it is already made at least hundreds of milliseconds (and even up to 10 seconds) before any awareness state occurs (D’Ostilio & Garraux, 2012; Haynes, 2011; Libet et al., 1983). Nevertheless, it has been pointed out that it is important to further explore the strength of the relation between unconscious predictive brain processes and subsequent decisions from a conscious perspective, in order to eliminate potential shortcomings of experiments claiming that awareness is not contributing to action selection or execution (Haynes, 2011).

Haggard and co-workers explored the role of conscious awareness in action selection and found that both predictive (prior to action execution) and inferential (after execution of the action) processes play a role in action awareness (Moore & Haggard, 2008). They stated:

“Our results suggest that both predictive and inferential processes contribute to the conscious awareness of operant action. The relative contribution of each of these processes seems to be context dependent. When we can predict the consequences of our actions, as in a high action-effect contingency block, the awareness of action reflects these predictions. This would provide us with a predictive sense of our own agency. In addition, our results show clear evidence that inferential processes also influence the conscious awareness of operant action. ... ... ... The interaction between predictive and inferential processes is of particular interest. ... ... ... The time course over which information about action is built up may be an important clue to this interaction. ... ... ... Sensory feedback provides more precise evidence about actions and their effects. This evidence becomes available only after a short sensory delay, but can then be transferred to memory. Thus, reliable and enduring sensory evidence replaces short-lived predictive estimates. We suggest that awareness of action therefore switches from a predictive to an inferential source as the action itself occurs, and as sensory information becomes available.” (Moore & Haggard, 2008, pp. 142–143)

Haggard and co-workers showed that awareness of an action is a dynamic combination of both prior awareness (i.e., awareness of action effect prediction) and retrospective awareness (i.e., awareness of the effects of an action) through
predictive motor control and inferential sense-making relative to action execution (Haggard, Clark, & Kalogeras, 2002; Haggard & Clark, 2003; Haggard, 2005, 2008). This strengthens the idea that awareness has a causal influence on decision making together with perceptual attraction and sensory consequences (prior to action execution). More importantly, awareness has a retrospective aspect, and this retrospective awareness is influenced by prior conscious awareness and affective interaction with the environment. Furthermore, Haggard and co-workers presented a new phenomenon called *intentional binding*: the time that is experienced between a person’s voluntary action and its perceived outcome is smaller when awareness was pre-existing compared to when it did not pre-exist (Haggard et al., 2002). This phenomenon has been argued to be the effect of either prior awareness or retrospective awareness, depending on different experimental setups. To investigate the relation with prior awareness, transcranial magnetic stimulation (TMS) was randomly applied over the motor cortex, in order to disrupt awareness. As a result, a significantly weakened intentional binding was observed, which provides evidence for the necessity of prior awareness (Moore & Obhi, 2012). Similarly, there have been experiments to analyse the influence of retrospective awareness. For example, by selecting some tasks where the prediction of the action outcome is difficult (or impossible) and offering a ‘tone’ after action execution (Moore & Haggard, 2008), it has been observed that retrospective processes play a role when prior predictive processes are absent (or when prediction was minimal). Therefore, by considering both prior and retrospective aspects, we can explain the impact of awareness in a broader manner.

In addition to the information provided above, the recent literature in cognitive neuroscience includes more results. In (Catherwood et al., 2014), evidence has been reported for Level 1 SA, Level 2 SA, and Level 3 SA. With all this information, it is clear that the current body of knowledge related to action selection may play an important role in further understanding the processes behind SA. Therefore, instead of merely believing that SA is only a product, these findings show the importance of interpreting SA both as a product and as a process.

9.3. Description of the Computational Model

The model in this paper is an extension of the previous works presented in (Thilakarathne, 2014a; Thilakarathne, to appear), by incorporating the idea of retrospective processes, together with the other evidence mentioned in Section 2. Figure 1 presents the cognitive model for SA and Table 1 summarizes the abbreviations used. The model takes inputs from two world states WS($s_i$) and WS($b_i$), where $s$ is a stimulus (that can be either external or internal to the agent) that may lead to an action execution, and $b_i$ represents the effects of the execution of an action $a_i$. The model accepts multiple inputs in parallel (it is also possible to have
a compound single input that triggers many options). Therefore, in this model, external input is a vector $s_k$, $k = 1, 2, \ldots$ where $k$ inputs are taken in parallel. The input world state $WS(s_k)$ leads to a sensor state $SS(s_k)$, and subsequently to a sensory representation state $SR(s_k)$. Moreover, the model includes both conscious and unconscious aspects. The following states are considered to be unconscious and contributing to bottom-up processes:

1. sensory representation of stimulus $s_k$ and effect $b_i$
2. performative desires for $b_i$
3. preparation for action $a_i$
4. feeling of $b_i$
5. perception of stimulus $s$ and effect $b_i$
6. prior ownership of action $a_i$ with effect $b_i$
7. retrospective ownership of action $a_i$ with effect $b_i$

In contrast, the following states represent more conscious influences, contributing to top-down processes:
subjective desires for effect bi
attention state for stimulus sk on effect bi
conscious intention state for stimulus sk on effect bi
prior-awareness state for action ai with effect bi and stimulus sk
retrospective-awareness state for action ai with effect bi and stimulus sk

Table 1: Nomenclature for Fig. 1

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WS(W)</td>
<td>world state W (W can be either stimulus s, or effect b)</td>
</tr>
<tr>
<td>SS(W)</td>
<td>sensor state for W</td>
</tr>
<tr>
<td>SR(W)</td>
<td>sensory representation of W</td>
</tr>
<tr>
<td>PD(bi)</td>
<td>performative desires for bi</td>
</tr>
<tr>
<td>SD(bi)</td>
<td>subjective desires for bi</td>
</tr>
<tr>
<td>PA(ai)</td>
<td>preparation for action ai</td>
</tr>
<tr>
<td>Per(bi,sk)</td>
<td>perception state for the effect of sk on bi</td>
</tr>
<tr>
<td>F(bi)</td>
<td>feeling for action ai and its effects bi</td>
</tr>
<tr>
<td>PO(ai, bi)</td>
<td>prior ownership state for action ai with bi</td>
</tr>
<tr>
<td>Att(bi,sk)</td>
<td>attention state for the effect of sk on bi</td>
</tr>
<tr>
<td>CInt(bi,sk)</td>
<td>conscious intention state for sk on bi</td>
</tr>
<tr>
<td>PAwr(ai,bi,sk)</td>
<td>prior-awareness state for action ai with bi and sk</td>
</tr>
<tr>
<td>EA(ai)</td>
<td>execution of action ai</td>
</tr>
<tr>
<td>RO(ai, bi)</td>
<td>retrospective ownership state for action ai with bi</td>
</tr>
<tr>
<td>RAwr(ai,bi,sk)</td>
<td>retrospective-awareness state for action ai with bi and sk</td>
</tr>
<tr>
<td>EO(ai,bi,sk)</td>
<td>communication of ownership and awareness of ai</td>
</tr>
</tbody>
</table>

- subjective desires for effect bi
- attention state for stimulus sk on effect bi
- conscious intention state for stimulus sk on effect bi
- prior-awareness state for action ai with effect bi and stimulus sk
- retrospective-awareness state for action ai with effect bi and stimulus sk

The dotted box in Fig. 1 represents the boundary between external and internal states: the states inside the dotted box represent internal cognitive states (in general those have a relatively higher firing rate), whereas all the other states are external.

The unconscious bottom-up process of action selection is modelled by combining Damasio’s as if body loop (cf. Damasio, 2010) and James’s body loop (cf. James, 1884).

- as-if body loop: PA(ai) → SR(bi) → F(bi)
- body loop: PA(ai) → EA(ai) → WS(bi) → SS(bi) → SR(bi) → F(bi)

According to Damasio, the cognitive process of action selection is based on an internal simulation process prior to the execution of an action. Effects of each relevant action option PA(ai) (a stimulus s will have many options i=1..n) are evaluated (without actually executing them) by comparing the feeling-related valuations associated to their individual effects. Each preparation state PA(ai) for action option ai suppresses preparation of all its complementary options PA(ai) with j≠i (see Fig. 1), and therefore by a kind of winner-takes-all principle, naturally the option that has the highest valued effect felt by the agent will execute through the body loop (through this agent does not follow a random selection mechanism but by
providing necessary weight values for each option in domain specific manner it can adhere to the environment). Furthermore, according to the literature, the predictive effect and sensed actual effect of the action are added to each other through an integration process (cf. Moore & Haggard, 2008); this is expected to be reflected in this model through the sensory representation of \( b_i \). Through this, it is also possible to demonstrate the difference between the case where both predictive and actual effects are the same and the case where they are not.

This process is further strengthened by embedding performative desires for \( b_i \) and perception states for \( s_k \) on \( b_i \). The performative desire state \( PD(b_i) \) facilitates short-term interests/goals that influence either selecting or rejecting an action due to its satisfactory or less satisfactory valuation (cf. Thilakarathne and Treur 2013a). Therefore, through this state, the agent has the ability to strengthen the current action selection based on its desires (through this, also some bias can be injected into the process). In parallel to the action preparation process, the perception state \( \text{Per}(b_i, s_k) \) also develops based on the salient features of the stimulus \( s_k \); this will strengthen the bottom-up process which leads to even further strengthening of the action preparation process (due to the activation that spreads from \( \text{Per}(b_i, s_k) \) to \( \text{PA}(a_i) \)). Furthermore; by having a suppressive link from the \( \text{Per}(b_i, s_k) \) state to itself (see Fig. 1), the competition among perceptual entities as mentioned in Section 2 is represented (cf. Lavie, 2005). Each suppressive link’s negative effect (strength) is relatively proportional to the strength of that particular perceptual state (i.e. \( \text{Per}(b_i, s_k) \)) and therefore the perception state for element \( i \) that has the highest activation suppresses its complements most strongly. As a result, the strongest candidate will dominate the competition and naturally will contribute to a stronger perceptual load. This suppression mechanism is in line with the lateral inhibition process (i.e., if a particular representation accumulates more evidence, that representation will suppress its fellow representations) (see Aron, 2007). This is an improvement over the model presented in (Thilakarathne 2014a), where biases were injected through different (negative) weight values for complementary links. Due to the developed perceptual load, the agent will not automatically attend to other salient features (unless a particular attention is put on some salient feature intentionally). Furthermore, this selected perception will be strengthened by the agent’s attention and subjective desires as well (see Fig. 1); this will be explained later.

While the agent is passively (unconsciously) performing an action selection as explained in Section 2 (for more details see Thilakarathne, 2014b), the agent starts to activate bottom-up attention (this is represented by the link from \( F(b_i) \) to \( \text{Att}(b_i, s_k) \)). The main functionality of the bottom-up attention is to pass current information into higher order cognitive states (i.e., states which are assumed to be closer towards the emergence of consciousness). Due to this bottom-up attention, the agent will activate its subjective desires of \( b_i \), which in turn leads to a conscious intention
of stimulus $s_k$ and effect $b_i$, and subsequently back to the attention state again. This cyclic process represents the transformation from bottom-up to top-down. Furthermore, this is in line with the idea of transforming Level 1 SA to Level 2 SA in Endsley’s model in terms of a dynamic process (i.e., from perception to comprehension) (Endsley, 1995; Endsley, 1999, 2000). Intention is considered to trigger goal directed preparation (see Thilakarathne, 2014b) and therefore this model includes an effect from the conscious intention state $\text{CInt}(b_i, s_k)$ to $\text{PA}(a_i)$. This link strengthens the option of action $a_i$ but suppresses its complementary options for all $\text{PA}(a_j)$ with $j \neq i$. This is also a part of the top-down process.

Once the attention (and its subjective aspects) has been developed, it injects conscious biases (through the top-down attention) into the action preparation and perception states. This is represented through the links from $\text{Att}(b_i, s_k)$ to $\text{PA}(a_i)$ and $\text{Per}(b_i, s_k)$, and these links (purple dotted arrows) play a special role: while activating the matching option (i.e. $i^{th}$ option) they suppress all complements of the $i^{th}$ option. The suppressive effects of this phenomenon are in line with the voluntary inhibition process (i.e., intentional suppression of an irrelevant response, stimulus, or memory) (see Aron, 2007). This emphasizes the conscious influence on action formation, and therefore attention will quickly enable the agent’s concentration, which may shorten the time required for action selection. More importantly, particular to the perceptual load, this will strengthen the current perception even further, and due to strong subjective feelings the agent may not be able to shift its attention easily (nevertheless, over longer time spans, attention will naturally get diluted; however, these effects are yet not included in this model).

Together with these processes, the agent will develop a state of ownership, which mainly determines to what extent an agent attributes an action to himself or to another agent. This particular aspect is important when it comes to situations where collaborative situation awareness plays a role, e.g. through collective decision making (although this is not in the scope of this paper). Also, as explained in previous works (see Thilakarathne, 2014a, 2014b), the agent will develop an awareness state of action $a_i$ that is related to effect $b_i$ and stimulus $s_k$. According to Harggard (Moore & Haggard, 2008; Haggard, 2008), there may be an influence from awareness states to action selection; therefore, this model includes a link from prior-awareness state $\text{PAwr}(a_i, b_i, s_k)$ to the action execution state $\text{EA}(a_i)$ (however, note that there are also claims that awareness of motor intentions does not have any influence on action execution, but emerges after action preparation and just before action execution (Libet et al., 1983; D’Ostilio & Garraux, 2012; Haynes, 2011)). Due to the empirical evidence that supports that awareness appears just before action execution, the current model includes that aspect by having awareness be affected mainly by higher order cognitive states (as per the processes mentioned in Section 2); also, it does not affect many other states directly. The agent will execute
the selected action $a_i$ (through the interplay between bottom-up and top-down processes) and then this action will have an effect in the environment (through world state $WS(b_i)$), and be sensed again, as explained earlier through the body loop.

Once an action is executed, this model also considers the retrospective aspects that Haggard and co-workers pointed out (Moore & Haggard, 2008; Haggard & Clark, 2003). First, the agent develops retrospective ownership, which is followed by retrospective awareness. The retrospective ownership state $RO(a_i, b_i)$ is affected by the states $PO(a_i, b_i)$, $F(b_i)$, and $EA(a_i)$. Once the state $RO(a_i, b_i)$ is developed, it has a suppressive effect on $PO(a_i, b_i)$ and this strengthens the cognitive shift from predictive to inferential processes (cf. Moore and Haggard 2008; Haggard and Clark 2003). The retrospective awareness state $RAwr(a_i, b_i, s_k)$ is affected by the prior awareness state $PAwr(a_i, b_i, s_k)$, as well as by $RO(a_i, b_i)$ and $F(b_i)$. Once the $RAwr(a_i, b_i, s_k)$ state is activated it also has a suppressive effect on $PAwr(a_i, b_i, s_k)$. In this model, when there is a mismatch between what is predicted and what is actual, this mismatch can be realised through prior and retrospective awareness states: when predicting a strong satisfactory effect, the agent develops strong prior awareness, but when not sensing the same predicted effect the agent will develop poor retrospective awareness of that predicted effect. Through this mechanism, the agent can examine the performance of an executed action relative to its predicted effects on a particular situation. Having such a mechanism to interpret a mismatch between predicted and occurred is important in many SA scenarios; therefore, this feature will be useful in complex simulations that include multiple episodes. More importantly, through this addition, the agent can reconsider possible alternative options by having realised that an action does not lead to a particular effect as expected. Finally, the agent has the ability to communicate the process through state $EO(a_i, b_i, s_k)$. The state $EO(a_i, b_i, s_k)$ is affected by the retrospective awareness state $RAwr(a_i, b_i, s_k)$, as well as by $RO(a_i, b_i)$ and $Clnt(b_i, s_k)$. Through this state it is assumed that other agents or systems can interact with this agent, in order to execute collective SA-based decisions (which are not in the scope of this paper).

### 9.4. Specification of Model Compilation

Each connection between states has been given a weight value (where $\omega_{ji}$ represents the weight of the connection from state j to i) that varies between 1 and -1 as indicated in Table 2. Weight values are non-negative in general, except if they represent a suppressive (or inhibiting) link (see caption of Fig. 1). To model the dynamics following the connections between the states as temporal–causal relations, a dynamical systems perspective is used, as explained in (Treur, 2013).

Furthermore, each state includes an additional parameter called speed factor $\gamma_i$, indicating the speed by which the activation level of the state ‘i’ is updated upon
Two different speed factor values are used, namely fast and slow values: fast values are used for internal states and slow values for external states (i.e., for WS(W), SS(W), EA(ai), and EO(ai, bi, sk)). The level of activation of a state depends on multiple other states that are directly attached to it. Therefore, incoming activation levels from other states are combined to some aggregated input and affect the current activation level according to differential equation (1). As the combination function for each state, a continuous logistic threshold function is used: see equation (2), where $\sigma$ is the steepness, and $\tau$ the threshold value. When the aggregated input is negative, equation (3) is used. To achieve the desired temporal behavior of each state as a dynamical system, the difference equation represented by equation (4) is used (where $\Delta t$ is the time step size).

\[
\frac{dy_i}{dt} = y_i \left[ f(\sigma, \tau, \sum_j \omega_{ij} y_j) - y_i \right] 
\]

\[
f(\sigma, \tau, X) = \left( \frac{1}{1 + e^{-\sigma (X - \tau)}} - \frac{1}{1 + e^{\sigma \tau}} \right) (1 + e^{\sigma X}) \; \text{when} \; X > 0
\]

\[
f(\sigma, \tau, X) = 0; \; \; \; \; \text{when} \; X \leq 0
\]
The complete formalisation of the cognitive model in computational form is provided below. For the dynamics of each Local Property (in Table 2 the column LP refers to (temporally) Local Properties), a formalisation in the LEADSTO language is used, which has been shown to be an appropriate approach to model dynamic behaviours of computational cognitive models (cf. Bosse, Jonker, Van Der Meij, & Treur, 2007). The local properties LP1 to LP18 specify the process of updating the activation level of the ‘to state’ based on the activation levels of the ‘from states’, based on the mentioned formalisation. Each state property has a strength represented by a real number between 0 and 1 through variables V (with subscripts) that run over these values. Each LP in rule-based LEADSTO notation and differential equation form can be found in the Appendix.

\[ y_i(t + \Delta t) = y_i(t) + \gamma_i \left[ f\left(\sigma, \tau, \sum_{j \in s(i)} \omega_{ij} y_j\right) - y_i(t)\right] \Delta t \quad (4) \]

9.5. Analysis of SA on the Proposed Model based on Simulations

In this section by simulations it will be explained how situation awareness related incidents can be explained through this proposed model. For this, three situations were selected from the document ‘Enhancing Situational Awareness2’ in ‘Flight Operations Briefing Notes’ from the Airbus Company. They have provided 3 generic examples for each of the three levels of the SA described by Endsley:

- For Level 1 SA: ‘Focusing on recapturing the LOC and not monitoring the G/S’
- For Level 2 SA: ‘Applying a fuel imbalance procedure without realizing it is an engine fuel leak’
- For Level 3 SA: ‘Expecting an approach on a particular runway after having received ATIS information and being surprised to be vectored for another runway’

In addition to these, a fourth simulation describes a situation where the agent prepared for an action by expecting a particular effect \(b_1\), though its actual effect after execution is different: \(b_2\). These four examples were simulated using an implementation in Java, based on the mathematical basis explained in the previous section. For each scenario, four different sets of input data were used in XML format with dedicated parameter values. Table 3 presents connection weight values and Table 4 presents threshold (\(\tau\)) and steepness (\(\sigma\)) values used in configurations of simulations of this cognitive model. Additionally, for:

- step size (Δt) value 0.25 is used
- fast speed factor (γ) value 0.9 is used
- slow speed factor (γ) value 0.6 is used

The main challenge in this approach is that there is no real detailed data set that can be compared to the output of the agent model, in order to estimate parameters. Only certain features of the behaviour of each cognitive state are known for different scenarios, based on neurological, behavioural and affective evidence from the literature (for example, prior awareness should occur before action execution and after prior ownership, and there should be a dip in the sensory representation in-between predictive representation and inferential representation, et cetera). To identify the parameter values, a systematic approach is used. For this approach it is a necessary condition to select multiple scenarios (minimum is 3 but having more will improve the quality of the results) which are interrelated from a functional point of view. For example, the first scenario is considered to be a reference scenario and the second scenario is different from that just by very few weight value changes (see Table 3: Scenario 1 and 2) and this should be a common feature in all scenarios selected for this approach. This interrelation among scenarios is very important for a minimum number of parameter changes, enabling us to identify a generic parameter value set for the model. In this parameter estimation approach the idea is as follows:

- The first scenario is addressed and parameter values are calibrated to simulate the behaviour as identified through the literature
- Then, by using the obtained parameter value set, by changing just a few (scenario-related) weight values it is checked whether the model with these parameter settings is able to generate the behaviour for the second scenario.
- If this provides a simulation with a pattern as expected (without changes to the previously obtained parameter values, except for the changes particular to the current scenario), then it provides good confidence on the currently identified parameter value set.
- But if not, then it is necessary to change the parameter values of the first simulation (based on the sensitivity of certain parameters on the required final output) until the behaviours for both simulations are satisfactory.
- This approach is incrementally extended to each scenario until a generic parameter value set for all the scenarios has been identified. For any new scenario, if any changes to the previously obtained parameter values are required, then all previously addressed scenarios are readdressed.
- In the first few iterations it may be challenging to identify a parameter value set, but over time the convergence is really fast and the identified
parameter value set well become more generic as fewer changes will be required

Table 3: Connection weight values used for cognitive agent model. In here if a value of a particular weight is empty for a scenario that means it is equal to the value of that in the Scenario 1 Stimulus 1. Furthermore if a value is ‘–’ then such a link was not existed for that scenario and furthermore red colour ‘0.x,0.y’ presents that the particular link suppresses its mapping inputs and the remaining complement subsequently.

<table>
<thead>
<tr>
<th>To State</th>
<th>Weights</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
<th>Scenario 4</th>
<th>For effect b₁</th>
<th>For effect b₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS(sₙ)</td>
<td>ω₉₁</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR(sₙ)</td>
<td>ω₂₂</td>
<td>1.0</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PD(bₙ)</td>
<td>ω₃₀</td>
<td>0.8</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA(aᵢ)</td>
<td>ω₄₄</td>
<td>0.5</td>
<td>0.1</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₄₅</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₄₆</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₄₇</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₄₈</td>
<td>1.0</td>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₄₉</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₅₀</td>
<td>0.9</td>
<td></td>
<td>0.1</td>
<td>0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₅₁</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₅₂</td>
<td>0.9</td>
<td></td>
<td>0.1</td>
<td>0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₅₃</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per(bᵣ, sₙ)</td>
<td>ω₁₄</td>
<td>0.9</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₁₅</td>
<td>0.8</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₁₆</td>
<td>0.8</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₁₇</td>
<td>0.8</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₁₈</td>
<td>0.9</td>
<td></td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₁₉</td>
<td>0.9</td>
<td></td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₂₀</td>
<td>0.9</td>
<td></td>
<td></td>
<td>0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR(bᵣ)</td>
<td>ω₂₁</td>
<td>0.9</td>
<td></td>
<td>0.8</td>
<td>0.8</td>
<td>0.7</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>ω₂₂</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₂₃</td>
<td>0.9,0.9</td>
<td></td>
<td>0.7,0.1</td>
<td></td>
<td>0.0,0.9</td>
<td></td>
</tr>
<tr>
<td>F(bᵣ)</td>
<td>ω₂₄</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>ω₂₅</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td>0.8</td>
<td>0.0</td>
</tr>
<tr>
<td>Att(bᵣ, sₙ)</td>
<td>ω₂₆</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₂₇</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₂₈</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD(bᵣ)</td>
<td>ω₂₉</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₃₀</td>
<td>0.9</td>
<td></td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₃₁</td>
<td>0.9</td>
<td></td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ClInt(bᵣ, sₙ)</td>
<td>ω₃₂</td>
<td>0.2</td>
<td></td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₃₃</td>
<td>0.9</td>
<td></td>
<td></td>
<td>0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PO(aᵢ, bᵣ)</td>
<td>ω₃₄</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₃₅</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ω₃₆</td>
<td>0.4,0.9</td>
<td></td>
<td>0.4,0.1</td>
<td>0.9,0.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Threshold (τ) and Steepness (σ) values used in configurations of simulations. Note: For four states $F(b_i)$, $RO(a_i, b_i)$, $RAwr(a_i, b_i, s_k)$, and $EO(a_i, b_i, s_k)$ different threshold values used (in red colour) for the fourth scenario “effect $b_2^2$” in Table 3.

9.5.1. Simulation for a Level 1 SA example incident

According to the statistics provided by Endsley in the aviation domain (see Endsley & Garland, 2000; Endsley, 1999), 76% of the errors related to poor SA were Level 1 errors, which are due to a failure to correctly perceive information. Therefore, from the point of view of the empirical data, an interesting challenge is to understand the reasons for such poor perception (relative to the appropriate perception that should have been developed) in Level 1 SA, and the processes behind the development of this perception. The following factors have been identified as the main physical/environmental reasons for this problem: data not available, data hard to detect, failure to monitor/observe data, misperception of data, and memory loss,
etc. (Endsley & Garland, 2000; Endsley, 1999). Most of these factors are mainly due to poor design or limitations in user interfaces. Nevertheless, the focus of this simulation is to model the cognitive behavior as a process for Level 1 SA and the perceptual load theory used for this purpose (as presented in the Section 2).

The reason behind this example incident is poor SA due to a failure to monitor/observe data. A pilot has observed only one device (LOC) though he/she was supposed to take into consideration data from two devices (LOC and G/S). Due to these missing data (or not able to attending to read the data on both devices), the pilot has developed a strong perception related to action selection only based on LOC data, while his perception should have been for action selection in line with the integrated reading of LOC and G/S. Due to this incomplete input the appropriate perception was unable to develop and as a consequence of that the pilot has not developed the ‘right’ situation awareness, but instead of an awareness based on incomplete situation information. For the sake of simplicity of the simulation it is assumed that the current stimulus includes salient features of the LOC device but not strong data from the G/S device. From that stimulus agent will prepare for two action options PA($a_1$) and PA($a_2$) where action $a_1$ is based on information from the device LOC and action $a_2$ is based on information from both devices.

To simulate this example, two input stimuli ($s_1$ and $s_2$) have been used. One input triggers action selection based on reading the LOC data only (in here through $s_1$), and the other one triggers action selection based on reading the data from both LOC and G/S combined (in here through $s_2$). Details on all the input information and parameter values for each state can be found in Table 3 and 4.

In this simulation, the main cognitive state of interest is the perception state, and the process used to influence this is the perceptual load. Therefore, for all weight values in the model for all options (two options have the potential to trigger $a_1$ and $b_1$ & $a_2$ and $b_2$), identical values are used, except for the connection weights between:

- $SR(s_k)$ to Per($b_i$, $s_k$)
- $PD(b_j)$ to Per($b_i$, $s_k$)
- $SD(b_j)$ to Per($b_i$, $s_k$)
- $Att(b_i, s_k)$ to Per($b_i$, $s_k$)

These four links are all related to the perception state, and contribute to the process behind perceptual load. Through this difference in weight values, the impact of strong perception on the rest of the process is shown. In this particular simulation, the ‘complements-suppressive’ link from Per($b_i$, $s_k$) to Per($b_j$, $s_i$) (where $j\neq i$) has the same weight value (therefore no bias is introduced through this link, but as mentioned in Section 3, the suppression is only proportional to the strength of the particular state). The following execution trace is expected from the agent in this case:
– External stimuli $s_1$, $s_2$ will occur and trigger preparation of actions $a_1$ and $a_2$ separately.
– Based on the external stimuli, performative desires $b_1$ and $b_2$ are generated separately.
– In parallel, perception of effect $b_1$ through stimulus $s_1$ and perception of effect $b_2$ through stimulus $s_2$ is generated.
– With the effects of strong perceptual load the perception of effect $b_1$ dominates while perception of effect $b_2$ does not have any chance to get developed.
– With the strong biased perception on $b_1$, through the perceptual load the other activations related to option $b_2$ do not get any chance to compete with option $b_1$.
– Based on the preparation state for $a_1$ and $a_2$, the sensory representation of predicted effect $b_1$ of $a_1$ and $b_2$ of $a_2$ are also generated.
– Nevertheless, the agent will show strong effects on option $a_1$ and the rate of activation for preparation of $a_2$ will quickly slow down and disappear subsequently.
– Based on this positive predicted effect and the other states for $a_1$, the agent acquires a positive predictive feeling for effect $b_1$.
– Based on sufficiently strong activations of the sensory representation, feeling and other states related to effect $b_1$, attention for effect $b_1$ is generated.
– The attention state for effect $b_1$ is followed by subjective desires for effect $b_1$ and subsequently a conscious intention for effect $b_1$.
– Based on the positive predicted effect and the other states related to $a_1$, a prior ownership state for action $a_1$ is generated.
– Prior ownership for action $a_1$ is followed by prior awareness, which is generated just before execution of the action.
– The agent will develop neither prior ownership nor awareness for option $a_2$.
– The prior ownership and prior awareness states for action $a_1$ lead to actual execution of action $a_1$.
– The execution of $a_1$ affects $b_1$ in a positive manner and, via sensing, also affect the sensory representation of $b_1$ and the feeling of $b_1$.
– At the same time, the sensory representation of $b_1$ is suppressed, due to its prior self ownership state.
– Based on the generated states, after the execution of action $a_1$, the agent develops a retrospective ownership state for action $a_1$.
– The retrospective ownership of action $a_1$ is followed by retrospective awareness of action $a_1$.
– Finally, the agent communicates this ownership and awareness related to option $a_1$.

Figure 2 provides the simulation results for this scenario (here, first graph is for the results of option 1 and second graph presents the results for option 2). Upon
receiving the two input stimuli, the agent will prepare for two action options \( PA(a_1) \) and \( PA(a_2) \), where action \( a_1 \) is based on information from the LOC device and action \( a_2 \) is based on information from both devices. From these graphs, it can be seen that the provided input stimuli have relatively large effects on \( SR(s_i) \) for both options (i.e., for \( i\in\{1,2\} \)), with the maximum of 0.53 per each. The states \( PD(b_i) \) also having strong activations at the beginning. Nevertheless, the agent only generates a strong action preparation state for action option \( a_1 \): the level of \( PA(a_1) \) becomes very high (with a max of 0.95), just like that of perception state \( Per(b_1, s_1) \) (with a max of 0.86). Instead, for action option \( a_2 \) it has a very weak \( Per(b_2, s_2) \) (max of 0.04) that contributes to the development of a poor \( PA(a_2) \) (max of 0.12). Hence, merely through this effect of strong perception (as Endsley highlighted), the agent has not developed the right situation awareness (in this case the ‘correct’ awareness would have been \( PAwr(a_2, b_2, s_2) \)). Instead, the ‘incorrect’ awareness state \( PAwr(a_1, b_1, s_1) \) (max of 0.70) is generated, based on wrong perception. Note, by the way, that SA is a subjective term and one can of course have discussion about when an SA state is appropriate for a given situation (however, this is beyond the scope of this paper). More importantly, this simulation has illustrated that the model can reproduce the effect of perceptual load under early selection while having identical weights for each option, except the four modifications mentioned (moreover, late selection is related to attention, and is considered to be outside the scope of this paper as well). Therefore, a bias has been injected through the process from perception to action selection process (from the link \( Per(b_i, s_i) \) to \( PA(a_i) \)), and mainly with the support of unconscious processes, the agent has moved towards action selection (note that Level 2 and 3 SA are assumed to be more conscious than Level 1 SA). Subsequently, the agent generates sufficient activation levels for all the other states related to option \( a_1 \), and finally executes the action \( EA(a_1) \) (max of 0.82) with a \( PAwr(a_1, b_1, s_1) \) of max 0.70. The maximum activation levels of the other states related to option 1 are: \( SR(b_1) \) is 0.5 (through predictive) 0.87 (through sensing), \( F(b_1) \) is 0.52 (through predictive) 0.81 (through sensing), \( Att(b_1, s_1) \) is 0.93, \( SD(b_1) \) is 0.86, \( Clnt(b_1, s_1) \) is 0.83, \( PO(a_1, b_1) \) is 0.62, \( RO(a_1, b_1) \) is 0.83, \( RAwr(a_1, b_1, s_1) \) is 0.82, and \( EO(a_1, b_1, s_1) \) is 0.88. This pattern is as expected based on previous works (see Thilakarathne, 2014b; Thilakarathne, to appear). Also, the agent has properly integrated the predictive effects and sensed actual effects through the sensory representation and feeling states. This can be explained by the two-step sigmoid curve (for \( SR(b_1) \) there is a slight saturation at time point 35, and then with the execution of \( EA(a_1) \) this is increased again with a higher steepness; the same behaviour can be found in the feeling state of effect \( b_1 \)) (cf. Moore & Haggard, 2008). If the agent’s predicted effect and its sensed actual effect would not be the same, then there would not be such two-step sigmoid behaviour.
9.5.2. Simulation for a Level 2 SA example incident

In this situation the problem is with the Level 2 SA (i.e., failure to comprehend the situation) and according to the incident the reason may be due to an incorrect...
mental model. In this situation the pilot has observed all the necessary data with a
correct and complete perception, and noted a problem with fuel usage. Nevertheless, the pilot was unable to realize that the reason was a fuel leak in the
engine, and therefore he has decided to follow fuel imbalance procedure, whereas
the recommendation is not to apply fuel imbalance procedure if fuel leak is
suspected (Cramoisi, 2010). The following execution trace is expected from the
agent in this case:

- External stimuli $s_1$, $s_2$ will occur and trigger preparation of actions $a_1$ (i.e., to execute the fuel imbalance procedure) and $a_2$ (i.e., to deal with a fuel leak in the engine) separately.
- Based on the external stimuli, performative desires $b_1$ and $b_2$ are generated separately.
- In parallel, perception of effect $b_1$ through stimulus $s_1$ and perception of effect $b_2$ through stimulus $s_2$ is generated.
- Based on the preparation state for $a_1$ and $a_2$, the sensory representation of predicted effect $b_1$ of $a_1$ and $b_2$ of $a_2$ are also generated.
- Based on these positive predicted effects and the other states for $a_1$ and $a_2$, the agent acquires positive predictive feelings for effects $b_1$ and $b_2$ separately.
- Based on sufficiently strong activations of the sensory representation, feeling and other states related to effect $b_1$ and $b_2$, the agent will develop bottom-up attention for effect $b_1$ and $b_2$ separately.
- The attention state for effect $b_1$ is followed by subjective desires for effect $b_1$; however, for effect $b_2$ no strong subjective desires will be generated.
- The subjective desire for effect $b_1$ is followed by the conscious intention for effect $b_1$; nevertheless, due to the lack of subjective desires for $b_2$ the agent is not generating a conscious intention for effect $b_2$.
- Therefore, the agent will show strong effects on option $a_1$ and the activation rate for preparation for $a_2$ will quickly slow down and subsequently disappear, due to the low activation of effect $b_2$.
- Based on the positive predicted effect and the other states for $a_1$, a prior ownership state for action $a_1$ is generated.
- The prior ownership for action $a_1$ is followed by prior awareness, which is generated just before execution of the action.
- The agent will develop neither prior ownership nor awareness for option $a_2$.
- The prior ownership and prior awareness states for action $a_1$ lead to actual execution of action $a_1$.
- The execution of $a_1$ affects $b_1$ in a positive manner and, via sensing, the sensory representation of $b_1$ and the feeling of $b_1$.
- At the same time, the sensory representation of $b_1$ is suppressed, due to its prior self ownership state.
Based on the generated states, after the execution of action $a_1$, the agent develops a retrospective ownership state for action $a_1$.

- The retrospective ownership of action $a_1$ is followed by retrospective awareness of action $a_1$.
- Finally, the agent communicates this ownership and awareness related to option $a_1$.

Figure 3 provides the simulation information for this scenario and it is in line with the expectations. Here for the given stimulus the agent will internally prepare for two action options: $a_1$ is to execute the fuel imbalance procedure and $a_2$ is to deal with a fuel leak in the engine. For this simulation all the states involve identical parameter values for the action options 1 & 2 separately, except for SD($b_1$) and Clnt($b_1$, $s_1$) together with few suppressive links (see Table 3). This shows the impact of subjective desires and intention of top-down control on other states. This is in line with the idea of transforming Level 1 SA to Level 2 SA in Endsley’s model (i.e., from perception to comprehension) and having wrong or poor compression skills on the right option/decision leads to this poor SA (Endsley, 1999, 2000). Nevertheless, correctness of SA is always relative and its quality can only be analysed well when a task is performed with an expert critiquing as a benchmark. Therefore, in this example it is assumed that the agent does not know that the action option $a_1$ leads to a poor SA state, and also immediately after executing action $a_1$ it does not show evidence to realise that the predicted option is invalid (the situation that the agent realises that there is a mismatch between predicted and sensed states will be simulated in scenario 4).

The agent starts action formation with the input stimulus that triggers two action options as mentioned. At the beginning (in first 25 time slots) it clearly shows that the rate of activation for Per($b_1$, $s_1$) and Per($b_2$, $s_2$) are almost the same (similarly the other pairs: PA($a_1$), SR($b_1$), and F($b_1$)), but the development of SD($b_1$) and Clnt($b_1$, $s_1$), the rates of increase related to action option $a_2$ have been significantly declined. The states SD($b_2$) and Clnt($b_2$, $s_2$) have not been activated with sufficient strength (which was assumed to be the relevant mental model to interpret the situation as an engine fuel leak) and therefore the state Att($b_1$, $s_1$) has increased rapidly (with a max of 0.94) due to the cyclic dependency highlighted among SD($b_1$), Clnt($b_1$, $s_1$) and Att($b_1$, $s_1$). Therefore, naturally the agent has been led to select option $a_1$. Agent developed a strong prior awareness PAwr($a_1$, $b_1$, $s_1$) (with a max of 0.72) and has executed option $a_1$ (i.e., EA($a_1$)) with a maximum activation value of 0.84.

By using the same parameter values for each state of the respective action options, but only using different values for the states SD($b_1$) and Clnt($b_1$, $s_1$) (together with a few suppressive links), we have explained the behaviour of SA in Level 2: the inability of binding perceptual information relevant to the agent’s subjective goals through comprehension. As the agent has developed strong perception of both
effects, this will support the argument that the agent has properly perceived the information to understand that there is a problem with fuel usage. Nevertheless, by not having subjective desires and conscious intentions about a fuel leak in the engine (which is represented through $b_2$), the agent has not properly integrated the information into the correct mental model. This results in the execution of action $a_1$.

**Fig. 3:** Simulation details for Level 2 SA example. First graph presents the behavior of when dealing with fuel imbalance procedure, and the next graph represents the behavior of fuel leak procedure.
which represents the execution of the fuel imbalance procedure. This simulation is in line with what is claimed by Endsley as well as the neurocognitive evidence presented in Section 2. Furthermore, the agent has also developed retrospective effects in line with its predictive behaviour. This resulted in the action to communicate its decisions with a strong ownership and awareness (with a max of 0.9). Furthermore, a considerable improvement can be observed in this simulation when comparing it with the results of the same simulation in (Thilakarathne, 2014a). In particular, the effects of option $a_2$ do not continue to develop with the effects of option $a_1$ in this improved version.

**9.5.3. Simulation for a Level 3 SA example incident**

Scenario 3 is different from the previous two as according to this the pilot was expecting an approach on a particular runway (let’s say R14) and while he is preparing for that he gets an instruction from the air traffic controller (ATC) to be vectored for a different runway (let’s say R35). In this scenario it is assumed that landing on R14 is the most common action and therefore without getting a direct request from ATC the pilot was preparing for the habitual task. Due to this new ATC instruction now the pilot may be unable to immediately adjust for this new situation as he may have not load the necessary mental model to execute the new instruction. This may go together with the effect of ‘over-projection of current trends’ as mentioned in Section 1 as one of the possible reasons behind poor Level 3 SA. Therefore, it is assumed here that due to this over-projection of current trends, the pilot is unable to immediately project the necessary future actions. Therefore first s/he needs to internally suppress her/his current action execution and needs to get ready for the relevant action choice for the new ATC instruction. Simulated behavior of this situation is presented in Fig. 4. Two stimuli were used for this scenario but they occur at different time points: one at time $t=0$ and the other one at time $t=100$. More specifically, it has been assumed that at $t=100$ the agent is getting the ATC instruction and by that time the agent was already performing an action with the intention of approaching to R14 (labeled as action option $a_1$, whereas the new action after $t=100$ is labeled as $a_2$). The following execution trace is expected from the agent in this case:

- At time point 0, external stimulus $s_1$ will occur and trigger preparation of action $a_1$.
- Based on the external stimulus, performative desire $b_1$ is generated.
- In parallel, perception of effect $b_1$ through stimulus $s_1$ is generated.
- Based on the preparation state for $a_1$, the sensory representation of the predicted effect $b_1$ of $a_1$ is generated.
- Based on this positive predicted effect and the other states for $a_1$, the agent generates a positive predictive feeling for effect $b_1$. 
– Based on sufficiently strong activations of the sensory representation, feeling and other states related to effect $b_1$, attention is generated for effect $b_1$.
– The attention state for effect $b_1$ is followed by subjective desires for effect $b_1$ and subsequently a conscious intention for effect $b_1$.
– Based on the positive predicted effect and the other states for $a_1$, a prior ownership state for action $a_1$ is generated.
– The prior ownership for action $a_1$ is followed by prior awareness, which is generated just before execution of the action.
– The prior ownership and prior awareness states for action $a_1$ lead to the actual execution of action $a_1$.
– The execution of $a_1$ affects $b_1$ in a positive manner and, via sensing, affects the sensory representation of $b_1$ and the feeling of $b_1$.
– At the same time, the sensory representation of $b_1$ is suppressed, due to its prior self ownership state.
– Based on the generated states, after the execution of action $a_1$, the agent develops a retrospective ownership state for action $a_1$.
– The retrospective ownership of action $a_1$ is followed by retrospective awareness of action $a_1$.
– The agent communicates this ownership and awareness related to option $a_1$.
– At time point 100, the agent receives a new external stimulus $s_2$, which triggers preparation of action $a_2$.
– Based on the external stimuli, performative desire $b_2$ is generated.
– In parallel, perception of effect $b_2$ through stimulus $s_2$ is generated.
– Based on the preparation state for $a_2$, the sensory representation of the predicted effect $b_2$ of $a_2$ is generated.
– Based on this positive predicted effect and the other states for $a_2$, the agent generates a positive predictive feeling for effect $b_2$.
– Based on sufficiently strong activations of the sensory representation, feeling and other states related to effect $b_2$, attention is generated for effect $b_2$.
– The attention state for effect $b_2$ is followed by subjective desires for effect $b_2$ and subsequently a conscious intention for effect $b_2$.
– Due to the activations of action preparation, perception, attention, and conscious intention for effect $b_2$, cognitive control processes will be developed, which lead to suppression of the states related to effect $b_1$ for option $a_1$.
– While cognitive controlling is in action, based on the positive predicted effect and the other states for $a_2$, a prior ownership state for action $a_2$ is generated.
– The prior ownership for action $a_2$ is followed by prior awareness, which is generated just before execution of the action.
– This prior awareness also strengthens the cognitive control of option $a_1$. 
— The prior ownership and prior awareness states for action $a_2$ lead to the actual execution of action $a_2$.
— The execution of $a_2$ affects $b_2$ in a positive manner and, via sensing, affects the sensory representation of $b_2$ and the feeling of $b_2$.
— At the same time, the sensory representation of $b_2$ is suppressed, due to its prior self ownership state.
— Based on the generated states, after the execution of action $a_2$, the agent develops a retrospective ownership state for action $a_2$.
— The retrospective ownership of action $a_2$ is followed by retrospective awareness of action $a_2$.
— Finally, the agent communicates this ownership and awareness related to option $a_2$. Subsequently, the effects related to option $a_1$ are stopped completely, since the agent has successfully committed to option $a_2$.

From Figure 4 (here also two graphs have been provided merely for the clarity of images but information of both graphs obtained from a single execution) it shows that the agent has initiated action formation for option $a_1$ and has developed sufficiently high activation of PD($b_1$) (max of 0.56), PA($a_1$) (max of 0.95), Per($b_1, s_1$) (max of 0.86), SR($b_1$) (max of 0.50: predictive, 0.88: sensed), F($b_1$) (max of 0.60: predictive, 0.82: sensed), Att($b_1, s_1$) (max of 0.93), SD($b_1$) (max of 0.86), CInt($b_1, s_1$) (max of 0.84), PO($a_1, b_1$) (max of 0.62), PAwr($a_1, b_1, s_1$) (max of 0.70), EA($a_1$) (max of 0.82), RO($a_1, b_1$) (max of 0.83), RAwr($a_1, b_1, s_1$) (max of 0.82), and EO($a_1, b_1, s_1$) (max of 0.88) (in the order mentioned here). Nevertheless, having a new instruction at $t=100$, the agent has started to suspend its current action and enabling the relevant states to execute option $a_2$. Related to option $a_2$, the respective states have obtained slightly higher activation values in the same order as for option $a_1$. Furthermore, it can be clearly observed that to fully execute action $a_1$, the agent has taken roughly 60 time intervals but for $a_2$ to be fully activated it has taken more than 80 time intervals (due to the mental overload: to suppress the current action and to form the new action selection).

As shown in Table 3, identical weight values were used for all options, except for the suppressive links initiated from the action preparation, perception, attention and conscious intention states. These suppressive links were selected such that those related to option $a_1$ have small negative values while those related to option $a_2$ have large negative values. Therefore, merely due to this (unconscious and conscious) suppressive competition, the agent is able to gradually generate the necessary states for option $a_2$, while completely extinguishing all the states related to option $a_1$. This behaviour is in line with the information in (Thilakarathne, 2014b), which emphasises the interplay between bottom-up and top-down processes. Once a strong attention for effect $b_2$ is developed, top-down effects on the preparation for option $a_1$ take place. This strong attentional bias inhibits the activation of option $a_1$ and
once the prior awareness on effect $b_2$ is developed it further contributes to suppress the strong perception developed for $b_1$. Hence, this mechanism simulates the process where the pilot changes his expectation to land on runway R14 to the expectation to land on R35 with the use of cognitive control.

**Fig. 4:** Simulation details for Level 3 SA example. First graph presents the behavior of landing on runway R14 from time at 0, and the next graph represents the behavior of landing on runway R35 from time at 100.
9.5.4. Simulation for mismatch between the predicted and actual effects of an action

In most of the previous scenarios, the predicted effect positively affected the agent’s action execution; however, in reality that is not always the case. In some situations, one’s prior expectation may conflict with the retrospective effects of an action after its execution. This particular phenomenon is important from a learning perspective, especially in critical domains. In this final scenario, this phenomenon will be simulated. For this simulation a single input is considered, but it triggers two effects: the first one is $b_1$ (let’s assume this refers to a decreased altitude), which is the predicted effect of action $a_1$, whereas the second effect, $b_2$ (an increased altitude) is not predicted. Instead, this second effect $b_2$ is what actually will happen after execution of $a_1$. In Table 3, the relevant weights have been highlighted for each effect separately. For effect $b_1$, weight $\omega_{45}$ (i.e., from EA($a_1$) to WS($b_1$)) was set to ‘0’, as the effect the of actual execution of $a_1$ is not $b_1$; instead, the same weight for effect $b_2$ was set to ‘1.0’ to simulate the non-predicted effect of the action. The weight of the connection from WS($b_2$) to SS($b_2$) (i.e., $\omega_{46}$ for the second effect) was set to 1.0 to facilitate the effect of actual sensing. Also, according to this scenario, at the beginning the agent will only predict effect $b_1$ of action $a_1$, but not $b_2$. To realise this, the weight $\omega_{21}$ (i.e. PA($a_1$) to SR($b_2$)) was set to ‘0’ (the same explanation applies for weight $\omega_{23}$). Moreover, the weight $\omega_{25}$ from PD($b_1$) to F($b_2$) was set to ‘0’ (as there is no predictive effect on effect $b_2$ it is not possible to have a PD($b_2$)). The weight $\omega_{47}$ from EA($a_1$) to RO($a_1$, $b_1$) was set to ‘0’, as there will not be any retrospective ownership for action $a_1$. Furthermore, the weight $\omega_{49}$ for effect $b_2$ (i.e., from PO($a_1$, $b_2$) to RO($a_1$, $b_2$)) was set to ‘0’ as there is no prior ownership on this effect. Also as there is no prior awareness related to $b_2$, weight $\omega_{37}$ (i.e., from PAwr($a_1$, $b_2$) to RAwr($a_1$, $b_2$)) was also set to ‘0’. As there is no conscious intention for $b_2$ to perform action $a_1$, the weight of the link from CInt($a_1$, $b_2$) to EO($a_1$, $b_2$, $s_1$) was set to ‘0’. The following execution trace is expected from the agent:

- External stimulus $s_1$ occurs and triggers preparation of action $a_1$.
- Based on the external stimulus, performative desire $b_1$ is generated.
- In parallel, perception of effect $b_1$ of $a_1$ is generated.
- Based on the preparation state for $a_1$, a sensory representation of the predicted (positive) effect $b_1$ of $a_1$ is generated.
- Based on this positive predicted effect and the other states for $b_1$, the agent generates a positive predictive feeling for effect $b_1$.
- Based on sufficiently strong activations of the sensory representation, feeling and other states related to effect $b_1$, attention is generated for effect $b_1$.
- The attention state for effect $b_1$ is followed by the subjective desires for effect $b_1$ and subsequently a conscious intention for effect $b_1$. 
Based on the positive predicted effect and the other states for $a_1$, a prior ownership state for $b_1$ of $a_1$ is generated.

The prior ownership for action $a_1$ is followed by prior awareness, which is generated just before execution of the action.

The prior ownership and prior awareness states lead to the actual execution of action $a_1$.

The execution of $a_1$ does not affect $b_1$, but has a different effect, $b_2$. Therefore, by sensing this effect, the agent will develop a sensory representation of $b_2$ and a feeling of $b_2$.

Based on the generated states, after the execution of action $a_1$, the agent may develop a low retrospective ownership state for action $a_1$ with effect $b_2$, due to the conflict between the predicted effect and the effect that is actually sensed.

The retrospective ownership of action $a_1$ with effect $b_2$ is followed by retrospective awareness of action $a_1$ with effect $b_2$.

Finally, the agent communicates this ownership and awareness.

A simulation run of this scenario is shown in Figure 5. Indeed, this simulation follows the expected pattern described in the trace. The agent generates performative desires for $b_1$ at time point 3 (with a peak value of 0.47) together with a preparation state for action $a_1$ (with a peak value of 0.90). Based on these states, perception of effect $b_1$ of action $a_1$ is generated (with a peak value of 0.81). Based on this, a sensory representation of the predicted effect $b_1$ of $a_1$ is generated (with a peak value of 0.40), followed by a feeling of $b_1$ (with a peak value of 0.40). Subsequently, attention for $b_1$ is generated (with a peak value of 0.83), and is followed by subjective desires for $b_1$ (with a peak value of 0.80) and a conscious intention for $b_1$ (with a peak value of 0.69), respectively. These states contribute to the generation of prior ownership for $b_1$ (with a peak value of 0.43) and are followed by a prior awareness state for $b_1$ (with a peak value of 0.59). Hence, the agent has properly predicted the effects of $b_1$. As a result of the prior awareness and ownership states, the agent initiates the actual execution of action $a_1$, which propagates its actual effects through the (external) body loop. However, the execution of action $a_1$ does not affect $b_1$, but it has a different effect, $b_2$. Therefore, the sensory representation $b_1$ of $a_1$ does not behave as would be expected by adding the sensed actual effect to the predicted effect, and this effect also does not propagate to the feeling of $b_1$, as pointed out earlier (c.f. Moore & Haggard, 2008; Voss et al., 2010). Furthermore, the new effect $b_2$ even suppresses the activation of the sensory representation and the feeling of $b_1$. Instead, the sensory representation and feeling for $b_2$ have obtained strong activations, that lead to relatively low retrospective ownership and awareness states (mainly due to the absence of influences of prior ownership and awareness states, respectively). In contrast, for $b_1$ the agent has not developed any positive retrospective ownership and awareness.
Note that there is a clear activation of the sensory representation and feeling of $b_1$, but there is no ‘two-step increasing behaviour’ as observed in all the previous scenarios for these two states. This observation is a key element to explain the conflict experienced by the agent (i.e., the pilot expected a decrease in altitude, but the sensed effects have not been integrated, as there actually is an increase in altitude (cf. Moore & Haggard, 2008; Voss et al., 2010)). Furthermore, this can also be observed by the fact that there is prior ownership and prior awareness for effect $b_1$, but retrospective ownership and awareness for a different effect, $b_2$. Hence, through this process, the agent internally feels the conflict and may use this information to take the next action. This is in line with the SA models proposed by Neisser, which are inspired by the perception/action cycle (Neisser, 1976). This approach treats SA both as a product and a process (cf. Adams et al., 1995; Durso & Dattel, 2004; Klein et al., 2007). Through this addition, the agent can always scrutinize the situation when there is a conflict, and can move on to take a corrective action or a new action.

**9.6. Discussion**

This paper has presented a neurologically inspired cognitive model and has provided simulation results for 4 incident examples where poor SA was expected as put forward by Endsley. The obtained results explain the different scenarios with acceptable level of details as explained in section 4. This model uses relatively new

![Fig. 4: Simulation details for mismatch between the predicted and actual effects of an action.](image)
body of literature related to human cognition, and especially on action selection to explain the SA. This model has some differences compared to what Endsley suggests; mainly, it moves away from the idea that there is a causal chain from ‘Situation Awareness’ to ‘Decision Making’ to ‘Performance Evaluation’. In the proposed model, these 3 aspects are still exist but are more aligned with the findings from a neuroscience perspective. More specifically, this research shows how models that were designed according to the earlier cognitive science tradition and often assume linear causal cascades from sensory input to behavioural output, can be refined and enriched by incorporating more recent evidence on actual brain processes in which cyclic processes play a major role. Such model refinement often leads to dynamic systems style models with cyclic causal cascades instead of linear ones, as is clearly shown here (see also Cramoisi, 2010).

The simulations presented in this paper cover all three SA levels that Endsley highlighted. The first simulation shows the cognitive workings behind poor Level 1 SA, which mainly occurs due to a failure to correctly perceive information. Furthermore, this study has investigated the cognitive mechanisms behind this failure to correctly perceive information, and found perceptual load to be a key process behind this. Under the effects of perceptual load, the simulation results show why and how it may happen that a pilot observes only one device (the LOC) although he/she was supposed to take data from two devices (LOC and G/S) into consideration. Due to the high perception state developed for the LOC data, the agent was not able to perceive the data from the G/S device. As a result, the decision making process is dominated by the perception of the LOC information and finally leads to the execution of a corresponding action.

In the second simulation, in contrast, the agent does develop the correct perception. According to Endsley, poor Level 2 SA occurs due to a failure to comprehend the situation. In the simulated scenario, although an adequate perception of the situation regarding fuel usage, the agent is not able to bind this to the proper cognitive model and generate an intention to act on a fuel leak in the engine. Instead, the agent has developed an incorrect awareness on aspects related to fuel imbalance, and subsequently executes the (inappropriate) fuel imbalance procedure. This simulation thus explains the basics behind poor Level 2 SA, and shows how binding of wrong mental models will give rise to this situation from a dynamic modelling perspective.

The third scenario addresses a case of failure to project a future situation properly. In this case, the agent is assumed to be prepared for a habitual task (landing on runway R14) and indeed starts performing this action. However, while the agent is doing this, he receives an ATC message asking to vector for a different runway (R35, which is assumed to be not the default runway). Therefore, this simulation shows the ‘cognitive control’ process where one needs to shift away
from an action one is currently executing: i.e., a shift from landing on R14 to landing on R35. This shift is achieved through the interplay between bottom-up and top-down processes only, and details of the simulation show how the receipt of new instructions affects the states related to landing on R14.

Finally, the fourth simulation demonstrates a conflict between what is predicted and what actually occurs. Here, the agent is doing an action with the expectation that this will result in a decrease of altitude. However, once it performs the action, the agent senses that the altitude is actually not decreasing, but increasing. These four simple examples illustrate the features of the propose model in different situations, and explain the dynamics of situation awareness in cognitive terms. More importantly, a common configuration was used for all four simulations, and only with minimal differences on certain parameter settings, as mentioned in Table 3.

The simulations presented within this paper are mainly related to the aviation domain. Nevertheless, the model can also be used in other domains where SA related cognitive aspects are important. In general, a main usefulness of a model like this is for simulations of complex socio-technical systems (cf. Addyman & French, 2012). Simulations have been proposed as one of the promising techniques to study and analyse the behaviour in complex systems (e.g., air traffic control, ship navigation, emergency services) that cannot (or are practically infeasible to) be scrutinized through real experiments (due to unacceptable risks, high cost, poor level of detail, issues with repeatability and redesigning, etc.). The importance and manageability of simulations are growing, due to the increased availability of cheap computational power. Furthermore, multi-agent based simulation techniques make simulations easier to prototype and design for systems with a considerably large number of entities (Bonabeau, 2002). Nevertheless, a key limitation in this area is the lack of realistic models to be used for each individual agent. There are rule-based agent models that can be used for this purpose, but they are far away from reality (from a dynamic perspective). Therefore, it is important to design models that can be used as agents that generate the necessary behaviour in a way that is more in line with its real world basis (especially cognitive models inspired by neurocognitive evidence).

Having noted that human cognition is a complex system and that agent-based simulations require aspects of human cognition, a cognitive model like this could be a useful contribution (Sun, 2008). In addition, cognitive architectures (e.g., ACT-R, CLARION, SOAR) may provide more features for this purpose, as well as functional capabilities though a generic well known architectural basis (Taatgen & Anderson, 2010). Nevertheless, more focus is given in this work on modelling the details of human situation awareness through relatively new insights in brain activities, rather than modelling this through existing architectural constraints. This
specific cognitive model also includes features of a hybrid cognitive architecture (with the features of both symbolic and connectionist systems) that comprises features of both single-module architecture and/or multi-module homogeneous architecture (cf. Sun, 2001). This model includes the aspects of perceptual load, predictive processes, inferential processes, cognitive controlling, unconscious bias, conscious bias, and top-down and bottom-up interplay. Therefore; the potential applications of this model are high. More importantly, the model is easily extendable: it can be extended, among others, with processes like emotional awareness (Thilakarathne & Treur, 2014), intentional inhibition (Thilakarathne and Treur 2013a), and much more (Thilakarathne, 2014a, 2014b; Thilakarathne & Treur, 2013b, 2015; Thilakarathne, to appear). Also, having incorporated retrospective processes as an extension of previous work (Thilakarathne, 2014a, to appear), this paper expands the scope of the model. This model is in line with the perception/action cycle proposed by Neisser (1976) and its implementations (Adams et al. 1995; Durso and Dattel 2004; Klein et al. 2007). Nevertheless, retrospective aspects related to awareness have not been adopted in the way explained by Neisser, but have been inspired by the relatively new ideas proposed by Haggard and co-workers (Haggard, 2005, 2008; Haggard & Clark, 2003; Haggard et al., 2002; J. Moore & Haggard, 2008). Finally, this model also includes the aspects highlighted by Endsley (Endsley, 1995, 1988, 1999, 2000; Endsley & Garland, 2000). Therefore, we can conclude that this model covers a considerable spectrum of behaviour.

This work can be further extended to explain more complex scenarios in the aviation domain and in other domains where SA is important. For example, departing from scenario 4, this model can also be used to simulate scenarios that involve a longer time period (for example, a complete continuous episode of SA-driven operations, like air traffic management involving pilots, ground forces, air traffic controllers, etc.). For such complex simulations it is necessary to include this model within a dedicated simulation engine (depending on the domain) and apply it only for the agents for which it is required. Finally, it may be interesting to adapt this model to reduce its complexity, in order to enhance the performance of the simulations.

Acknowledgments

I wish to thank Prof. Jan Treur and Dr. Tibor Bosse at VU University Amsterdam, for their great support and supervision in all the phases of this work. This work is part of the SESAR WP-E programme on long-term and innovative research in air traffic management. It is co-financed by Eurocontrol on behalf of the SESAR Joint Undertaking (SJU).
Appendix

LEADSTO is a hybrid modeling language (Bosse et al., 2007) in which a dynamic property or temporal causal relation a→b denotes that when a state property a (or conjunction thereof) occurs, then after a certain time delay, state property b will occur. The time delay defined in LEADSTO is taken as a uniform time step Δt here. Below, in addition to the rule-based LEADSTO notation, each local property is also shown in differential equation form.

**LP1: Sensor state for stimulus** $s_k$: SS($s_k$)

If the world state for stimulus $s_k$ has level $V_i$
then after time duration $Δt$ the sensory state for stimulus $s_k$ will have the level $V_2$:
$$V_2 + γ [ f(ω_1 V_i) − V_2 ] Δt$$

$$\frac{dSS(s_k)(t)}{dt} = γ [ f(ω_1 WS(s_k)(t)) − SS(s_k)(t) ]$$

**LP2: Sensory representation state of stimulus** $s_k$: SR($s_k$)

If the sensory state for stimulus $s_k$ has level $V_i$
then after time duration $Δt$ the state sensory representation for stimulus $s_k$ will have the level $V_2$:
$$V_2 + γ [ f(ω_2 V_i) − V_2 ] Δt$$

$$\frac{dSR(s_k)(t)}{dt} = γ [ f(ω_2 SS(s_k)(t)) − SR(s_k)(t) ]$$

**LP3: Performative desires state for effect** $b_i$: PD($b_i$)

If the state sensory representation for stimulus $s_k$ has level $V_i$
and the complements of attention state for stimulus $s_k$ on effect $b_i$ have level $V_2$
then after time duration $Δt$ the state performative desires for an effect $b_i$ will have the level $V_3$:
$$V_3 + γ [ f(ω_3 V_i, ω_4 V_2) − V_3 ] Δt$$

$$\frac{dPD(b_i)(t)}{dt} = γ [ f(ω_3 SR(s_k)(t), ω_4 Att′(b_i, s_k)(t)) − PD(b_i)(t) ]$$

**LP4: Preparation state for action** $a_i$: PA($a_i$)

If the state sensory representation for stimulus $s_k$ has
level $V_1$
and the performative desires state for an effect $b_i$ has
level $V_2$
and the feeling state for an effect $b_i$ has level $V_3$
and the perception state for stimulus $s_k$ and effect $b_i$ has
level $V_4$
and the conscious intention state for stimulus $s_k$ and effect $b_i$
has level $V_5$
and the complements of conscious intention state for
stimulus $s_k$ and effect $b_i$ have level $V_6$
and the attention state for stimulus $s_k$ and effect $b_i$ has
level $V_7$
and the complements of attention state for stimulus $s_k$ and
effect $b_i$ have level $V_8$
and the complements of preparation state for action $a_i$ have
level $V_9$
then after time duration $\Delta t$ the preparation state for action $a_i$ will have the level
$V_{10}$:
$V_{10} + \gamma \left[ f(\omega_5 V_1, \omega_6 V_2, \omega_7 V_3, \omega_8 V_4, \omega_9 V_5, \omega_{10} V_6, \omega_{11} V_7, \omega_{12} V_8, \omega_{13} V_9) - V_{10} \right] / \Delta t$

$$\frac{dPA(a_i)(t)}{dt} = \gamma \left[ f(\omega_5 SR(s_k)(t), \omega_6 PD(b_i)(t), \omega_7 F(b_i)(t), \omega_8 Per(b_i, s_k)(t), \omega_9 Clnt(b_i, s_k)(t), \omega_{10} Clnt'(b_i, s_k)(t), \omega_{11} Att(b_i, s_k)(t), \omega_{12} Att'(b_i, s_k)(t), \omega_{13} PA'(a_i)(t) ) - PA(a_i)(t) \right]$$

**LP5: Perception state for stimulus $s_k$ on effect $b_i$: $Per(b_i, s_k)$**
If the state sensory representation for stimulus $s_k$ has
level $V_1$
and the performative desires state for an effect $b_i$ has
level $V_2$
and the subjective desires state for an effect $b_i$ has level $V_3$
and the attention state for stimulus $s_k$ and effect $b_i$ has
level $V_4$
and the complements of attention state for stimulus $s_k$ and
effect $b_i$ have level $V_5$
and the prior awareness state for an action $a_i$ with effect $b_i$
and stimulus $s_k$ has level $V_6$
and the complements of perception state for stimulus $s_k$
and effect $b_i$ have level $V_7$
then after time duration $\Delta t$ the perception state for stimulus $s_k$ and effect $b_i$ will have the level $V_8$: 

$$V_8 + \gamma \left[ f(\omega_{14}V_1, \omega_{15}V_2, \omega_{16}V_3, \omega_{17}V_4, \omega_{18}V_5, \omega_{19}V_6, \omega_{20}V_7) - V_8 \right] \Delta t$$

$$\frac{dPer(b_i, s_k)(t)}{dt} = \gamma \left[ f \left( \begin{array}{c} \omega_{14}SR(s_k)(t), \omega_{15}PD(b_i)(t), \omega_{16}SD(b_i)(t), \\ \omega_{17}Att(b_i, s_k)(t), \omega_{18}Att'(b_i, s_k)(t), \\ \omega_{19}PAwr(a_i, b_i, s_k)(t), \omega_{20}Per'(b_i, s_k)(t) \end{array} \right) \\ - Per(b_i, s_k)(t) \right]$$

**LP6: Sensory representation state of effect $b_i$: SR($b_i$)**

If the preparation state for action $a_i$ has level $V_1$ and the sensor state for effect $b_i$ has level $V_2$ and the prior ownership of action $a_i$ for $b_i$ and $s_k$ has level $V_3$ then after time duration $\Delta t$ the state sensory representation for an effect $b_i$ will have the level $V_4$: 

$$V_4 + \gamma \left[ f(\omega_{21}V_1, \omega_{22}V_2, \omega_{23}V_3) - V_4 \right] \Delta t$$

$$\frac{dSR(b_i)(t)}{dt} = \gamma \left[ f(\omega_{21}PA(a_i)(t), \omega_{22}SS(b_i)(t), \omega_{23}PO(a_i, b_i)(t)) - SR(b_i)(t) \right]$$

**LP7: Feeling state for action $a_i$ and its effects $b_i$: F($b_i$)**

If the state sensory representation for stimulus $s_k$ has level $V_1$ and the performative desires state for an effect $b_i$ has level $V_2$ then after time duration $\Delta t$ the state feeling for an effect $b_i$ will have the level $V_3$: 

$$V_3 + \gamma \left[ f(\omega_{24}V_1, \omega_{25}V_2) - V_3 \right] \Delta t$$

$$\frac{dF(b_i)(t)}{dt} = \gamma \left[ f(\omega_{24}SR(b_i)(t), \omega_{25}PD(b_i)(t)) - F(b_i)(t) \right]$$

**LP8: Attention state for stimulus $s_k$ on effect $b_i$: Att($b_i, s_k$)**

If the state feeling for an effect $b_i$ has level $V_1$ and the conscious intention state for stimulus $s_k$ and effect $b_i$ has level $V_2$ and the prior awareness state for an action $a_i$ with effect $b_i$ and stimulus $s_k$ has level $V_3$
then after time duration $\Delta t$ the attention state for stimulus $s_k$
and effect $b_i$ will have the level $V_4$:

$$V_4 + \gamma \left[ f(\omega_{26}V_1, \omega_{27}V_2, \omega_{28}V_3) - V_4 \right] \Delta t$$

\[
\frac{d\text{Att}(b_i, s_k)(t)}{dt} = \gamma \left[ f \left( \omega_{26}F(b_i)(t), \omega_{27}C\text{Int}(b_i, s_k)(t), \omega_{28}\text{P\text{Awr}(a_i, b_i, s_k)}(t) \right) - \text{Att}(b_i, s_k)(t) \right]
\]

**LP9: Subjective desires state for effect $b_i$: SD($b_i$)**

If the state sensory representation for stimulus $s_k$ has level $V_1$
and the attention state for stimulus $s_k$ and effect $b_i$ has level $V_2$
and the complements of conscious intention state for stimulus $s_k$ and effect $b_i$ have level $V_3$
then after time duration $\Delta t$ the subjective desire state for effect $b_i$ will have the level $V_4$:

$$V_4 + \gamma \left[ f(\omega_{29}V_1, \omega_{30}V_2, \omega_{31}V_3) - V_4 \right] \Delta t$$

\[
\frac{d\text{SD}(b_i)(t)}{dt} = \gamma \left[ f \left( \omega_{29}\text{SR}(s_k)(t), \omega_{30}\text{Att}(b_i, s_k)(t), \omega_{31}\text{C\text{Int}(b_i, s_k)}(t) \right) - \text{SD}(b_i)(t) \right]
\]

**LP10: Conscious intention state for stimulus $s_k$ on effect $b_i$: C\text{Int}(b_i, s_k)**

If the state perception state for stimulus $s_k$ and effect $b_i$ has level $V_1$
and the subjective desires state for an effect $b_i$ has level $V_2$
then after time duration $\Delta t$ the conscious intention state for stimulus $s_k$ and effect $b_i$ will have the level $V_3$:

$$V_3 + \gamma \left[ f(\omega_{32}V_1, \omega_{33}V_2) - V_3 \right] \Delta t$$

\[
\frac{d\text{C\text{Int}(b_i, s_k)}(t)}{dt} = \gamma \left[ f \left( \omega_{32}\text{P\text{Er}(b_i, s_k)}(t), \omega_{33}\text{SD}(b_i)(t) \right) - \text{C\text{Int}(b_i, s_k)}(t) \right]
\]

**LP11: Prior ownership state for action $a_i$ with effect $b_i$: PO($a_i, b_i$)**

If the state feeling for an effect $b_i$ has level $V_1$
and the preparation state for action $a_i$ has level $V_2$
and the retrospective ownership state of action $a_i$ for effect $b_i$ has level $V_3$
then after time duration $\Delta t$ the prior ownership of action $a_i$ for effect $b_i$ will have the level $V_4$:

$$V_4 + \gamma \left[ f(\omega_{34}V_1, \omega_{35}V_2, \omega_{36}V_3) - V_4 \right] \Delta t$$

\[
\frac{d\text{PO(a_i, b_i)}(t)}{dt} = \gamma \left[ f \left( \omega_{34}\text{P\text{Er}(a_i, b_i, s_k)}(t), \omega_{35}\text{PO(a_i, b_i)(t)} \right) - \text{PO(a_i, b_i)}(t) \right]
\]
\[
\frac{dPO(a_i, b_i)(t)}{dt} = \gamma \left[ f \left( \frac{\omega_{34} F(b_i)(t), \omega_{35} PA(a_i)(t), }{\omega_{36} RO(a_i, b_i)(t)} \right) - PO(a_i, b_i)(t) \right]
\]

**LP12: Prior awareness state for action \( a_i \) with effect \( b_i \) and stimulus \( s_k \):**

\( PAwr(a_i, b_i, s_k) \)

If the state prior ownership of action \( a_i \) for effect \( b_i \) has level \( V_1 \)
and the state feeling for an effect \( b_i \) has level \( V_2 \)
and the attention state for stimulus \( s_k \) and effect \( b_i \) has level \( V_3 \)
and the conscious intention state for stimulus \( s_k \) and effect \( b_i \) has level \( V_4 \)
and the retrospective awareness state for an action \( a_i \) with effect \( b_i \) and stimulus \( s_k \) has level \( V_5 \)
then after time duration \( \Delta t \) the state prior awareness for an action \( a_i \) with effect \( b_i \) and stimulus \( s_k \) will have the level \( V_6 \):

\[
V_6 + \gamma \left[ f(\omega_{37} V_1, \omega_{38} V_2, \omega_{39} V_3, \omega_{40} V_4, \omega_{41} V_5) - V_6 \right] \Delta t
\]

\[
\frac{dPAwr(a_i, b_i, s_k)(t)}{dt} = \gamma \left[ f \left( \frac{\omega_{37} PO(a_i, b_i)(t), \omega_{38} F(b_i)(t), \omega_{39} Att(b_i, s_k)(t), }{\omega_{40} Clnt(b_i, s_k)(t), \omega_{41} RAwr(a_i, b_i, s_k)(t)} \right) - PAwr(a_i, b_i, s_k)(t) \right]
\]

**LP13: Execution of action \( a_i \):** \( EA(a_i) \)

If the preparation state for action \( a_i \) has level \( V_1 \)
and the state prior ownership of action \( a_i \) for effect \( b_i \) has level \( V_2 \)
and the state prior awareness for an action \( a_i \) with effect \( b_i \) and stimulus \( s_k \) has level \( V_3 \)
then after time duration \( \Delta t \) the execution of action \( a_i \) state will have the level \( V_4 \):

\[
V_4 + \gamma \left[ f(\omega_{42} V_1, \omega_{43} V_2, \omega_{44} V_3) - V_4 \right] \Delta t
\]

\[
\frac{dEA(a_i)(t)}{dt} = \gamma \left[ f \left( \frac{\omega_{42} PA(a_i)(t), \omega_{43} PO(a_i, b_i)(t), }{\omega_{44} PAwr(a_i, b_i, s_k)(t)} \right) - EA(a_i)(t) \right]
\]

**LP14: World state of effect \( b_i \):** \( WS(b_i) \)

If the execution of action \( a_i \) state has level \( V_1 \)
then after time duration $\Delta t$ the world state for effect $b_i$ will have the level $V_2$:

$$V_2 + \gamma \left[ f(\omega_{45}V_i) - V_2 \right] \Delta t$$

$$\frac{dWS(b_i)}{dt}(t) = \gamma \left[ f(\omega_{45}EA(a_i)(t)) - WS(b_i)(t) \right]$$

**LP15: Sensor state of effect $b_i$: SS($b_i$)**
If the world state for effect $b_i$ has level $V_i$
then after time duration $\Delta t$ the sensor state for effect $b_i$ will have the level $V_2$:

$$V_2 + \gamma \left[ f(\omega_{46}V_i) - V_2 \right] \Delta t$$

$$\frac{dSS(b_i)}{dt}(t) = \gamma \left[ f(\omega_{46}WS(b_i)(t)) - SS(b_i)(t) \right]$$

**LP16: Retrospective ownership state for action $a_i$ with effect $b_i$: RO($a_i, b_i$)**
If the execution of action $a_i$ state has level $V_i$
and the state feeling for an effect $b_i$ has level $V_2$
and the state prior ownership of action $a_i$ for effect $b_i$ has level $V_3$
then after time duration $\Delta t$ the retrospective ownership state of action $a_i$ for effect $b_i$ will have the level $V_4$:

$$V_4 + \gamma \left[ f(\omega_{47}V_i, \omega_{48}V_2, \omega_{49}V_3) - V_4 \right] \Delta t$$

$$\frac{dRO(a_i,b_i)}{dt}(t) = \gamma \left[ f \left( \omega_{47}EA(a_i)(t), \omega_{48}F(b_i)(t), \omega_{49}PO(a_i,b_i)(t) \right) - RO(a_i,b_i)(t) \right]$$

**LP17: Retrospective awareness state for action $a_i$ with effect $b_i$ and stimulus $s_k$: RAwr($a_i, b_i, s_k$)**
If the retrospective ownership state of action $a_i$ for effect $b_i$ has level $V_1$
and the state feeling for an effect $b_i$ has level $V_2$
and the state prior awareness for an action $a_i$ with effect $b_i$ and stimulus $s_k$ has level $V_3$
then after time duration $\Delta t$ the retrospective awareness state for an action $a_i$ with effect $b_i$ and stimulus $s_k$ will have the level $V_4$:

$$V_4 + \gamma \left[ f(\omega_{50}V_i, \omega_{51}V_2, \omega_{52}V_3) - V_4 \right] \Delta t$$

$$\frac{dRAwr(a_i,b_i,s_k)}{dt}(t) = \gamma \left[ f \left( \omega_{50}EA(a_i)(t), \omega_{51}F(b_i)(t), \omega_{52}PO(a_i,b_i)(t) \right) - RAwr(a_i,b_i,s_k)(t) \right]$$
\[ \frac{dRAwr(a_i, b_i, s_k)(t)}{dt} = \gamma \left[ f \left( \omega_{50}RO(a_i, b_i)(t), \omega_{51}F(b_i)(t), \omega_{52}PAwr(a_i, b_i, s_k)(t) \right) \right] - RAwr(a_i, b_i, s_k)(t) \]

LP18: Communication of ownership and awareness state of \( a_i \) with effect \( b_i \) and stimulus \( s_k \): \( EO(a_i, b_i, s_k) \)

If the retrospective ownership state of action \( a_i \) for effect \( b_i \) has level \( V_1 \) and the conscious intention state for stimulus \( s_k \) and effect \( b_i \) has level \( V_2 \) and the state retrospective awareness for an action \( a_i \) with effect \( b_i \) and stimulus \( s_k \) has level \( V_3 \) then after time duration \( \Delta t \) communication of ownership and awareness of \( a_i \) for \( b_i, c_k, \) and \( s_k \) will have the level \( V_4 \):

\[ V_4 + \gamma \left[ f(\omega_{53}V_1, \omega_{54}V_2, \omega_{55}V_3) - V_4 \right] \Delta t \]

\[ \frac{dEO(a_i, b_i, s_k)(t)}{dt} = \gamma \left[ f \left( \omega_{53}RO(a_i, b_i)(t), \omega_{54}Clnt(b_i, s_k)(t), \omega_{55}RAwr(a_i, b_i, s_k)(t) \right) \right] \]

- \( RO(a_i, b_i)(t) \)
- \( F(b_i)(t) \)
- \( PAwr(a_i, b_i, s_k)(t) \)
- \( Clnt(b_i, s_k)(t) \)
- \( RAwr(a_i, b_i, s_k)(t) \)

References


http://doi.org/10.1016/j.tins.2011.02.003

http://doi.org/10.1146/annurev.psych.093008.131126

http://doi.org/10.1073/pnas.082080899

http://doi.org/10.3389/fpsyg.2012.00063

http://doi.org/10.1142/S0218213007003357

http://doi.org/10.1177/0018720814537070


http://doi.org/10.3389/fnhum.2012.00265


http://doi.org/10.1518/001872095779049543


Thilakarathne, D. J. (2014b). Modelling Dynamics of Cognitive Control in Action Formation with Intention, Attention, and Awareness. In 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) (pp. 198–205). IEEE. http://doi.org/10.1109/WI-IAT.2014.168


Chapter 10

An Analytical Model for Mathematical Analysis of Smart Daily Energy Management for Air to Water Heat Pumps

Seyed Amin Tabatabaei, Dilhan J. Thilakarathne, Jan Treur

Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.
e-mail: s.tabatabaei@vu.nl, d.j.thilakarathne@vu.nl, j.treur@vu.nl

Abstract: Having a substantial and increasing energy demand for domestic heating world wide together with decreasing availability of fossil fuels, the use of renewable energy sources for heating are becoming important. Especially air to water heat pumps have been suggested as an alternative for domestic heating. Using them the main challenge is to be sensitive to the timing of both the outdoor air temperature and the indoor energy demands in order to minimize the energy usage. Such a heating system should work efficiently and economically, thereby also taking into account comfort for the residents, with minimal human intervention. This paper presents an analytical model that can be used to mathematically analyse how to use energy more efficiently with an air to water heat pump. Results of such an analysis have been compared with another, simulation-based approach which has provided more confidence in the model. The model can be integrated in a smart thermostat to obtain more autonomous behavior for indoor heating demands against the changes in outside air temperature over the time.

Keywords: temperature maintain energy demand; temperature increase energy demand; indoor temperature; outdoor temperature; heat pump.

---

1 This chapter was published as:
http://doi.org/10.1016/j.egypro.2014.06.072
The names of the authors are ordered alphabetically reflecting the comparable contribution of each author.
10.1. Introduction

For smart energy management it is a challenge to improve the efficiency of energy usage for given demands. Heating and cooling remain processes that consume a lot of energy and it is predicted that this will increase exponentially (for the European Union statistics, cf. [1]). According to [1], from the total energy consumption 50% is for heating and from that 43% is for domestic heating. Therefore, the impact of improving the efficiency of domestic heating even by a relative small quantity will save billions of euro per year as a country and at a global scale. Also heating is a major contributor to carbon dioxide (CO2) emissions, and having more economical energy management will naturally contribute to a healthy environment and a sustainable development [2]. In practically all European countries heating is a necessary need at least part of the year. More and more domestic heating systems are considered that allow the use of renewable energy, in contrast to gas-based and oil-based heating systems that fully depend on non-renewable energy. As an important alternative, often heat pumps are suggested, which take most of their energy (e.g., from 50% to 80%) from the heat available in the ambient air, water or soil, and for the rest use electrical energy to drive them. Moreover, their electrical energy usage can be based on renewable production of, for example, solar and wind energy. Also for heat pumps it is important to make smart decisions about their use. This paper will focus on how to analyse in a mathematical manner the energy usage for heating based on an air-to-water heat pump. The proposed analytical model will provide the information to minimize the energy consumption. Thermostats were introduced to control the temperature in houses more easily and productively [3]. Since the invention of a basic thermostat by Cornelius van Drebbel [3], it has been extended with many features and by now it has come with features like fully programmable and smart behaviors [4]. Nevertheless, as put forward in [4] programmable thermostats have not shown the expected results and even in the USA the given certification for energy saving was discontinued by the EnergyStar™ in 2009. Therefore, it is a challenge to explore how to make thermostats smarter and more context sensitive with limited human intervention. An important issue may be continuous parameter adjustment for optimal benefits and facilitation of the real-time sensitivity to environmental changes (both indoor and outdoor).

As one step to address the above mentioned goals, in this paper the focus is on developing an analytical model for mathematical analysis that can be integrated in a thermostat allowing it to autonomously enabling the heating energy usage to become more economical. The temperature of the environment changes in a certain pattern over time which can be used as heuristic knowledge. Making a real-time estimation of this is a basis to locally adjust the heating needs with the intuition of
the current situation in the environment together with the outlook of temperature behavior in near future.

Below, in Section 2 the theoretical basis and background are described, and in Section 3 the analytical model is described. Section 4 presents some results of the use of this analytical model, and Section 5 discusses the obtained results and the future perspectives.

10.2. Background and theoretical basis

In this section the basic mathematical concepts will be introduced; see Table 1 for an overview. Subsequently they will be briefly explained.

10.2.1. Degree-days

The concept degree-day \( dd \) has been introduced to approximate the analysis of energy consumption and energy performance of a building based on historical data (e.g., [5]). The number of degree-days is defined as the summation of individual deviations between the outdoor temperature \( T_{od}(t) \) and a given indoor temperature \( T_{id}(t) \) in each time step for a time interval from \( t1 \) to \( t2 \) in [6]. This can be expressed mathematically as:

\[
dd(t1,t2) = \int_{t1}^{t2} (T_{id}(t) - T_{od}(t)) dt
\]

when \( T_{id}(t) > T_{od}(t) \) for all \( t \) with \( t1 \leq t \leq t2 \) (1)

10.2.2. Daytime and nighttime outdoor air temperature

The outdoor air temperature typically shows a 24h periodic behavior (e.g., [7, Table 1. Overview of concepts.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPF</td>
<td>seasonal performance factor</td>
<td></td>
</tr>
<tr>
<td>( T_{od}(t) )</td>
<td>outdoor temperature at time t</td>
<td>°C</td>
</tr>
<tr>
<td>( T_{id}(t) )</td>
<td>indoor temperature at time t</td>
<td>°C</td>
</tr>
<tr>
<td>( T_w )</td>
<td>water temperature of the heating system</td>
<td>°C</td>
</tr>
<tr>
<td>( dd )</td>
<td>degree days</td>
<td>°C day</td>
</tr>
<tr>
<td>( tmed )</td>
<td>heating energy demand for maintaining a temperature</td>
<td>kWh</td>
</tr>
<tr>
<td>( tied )</td>
<td>heating energy demand for increasing a temperature</td>
<td>kWh</td>
</tr>
<tr>
<td>( ed )</td>
<td>total heating energy demand</td>
<td>kWh</td>
</tr>
<tr>
<td>( eu )</td>
<td>heat pump electrical energy use</td>
<td>kWh</td>
</tr>
<tr>
<td>( \varepsilon )</td>
<td>energy loss per degree day</td>
<td>kWh / °C day</td>
</tr>
<tr>
<td>( C )</td>
<td>capacity: energy needed per degree increase of temperature</td>
<td>kWh / °C</td>
</tr>
</tbody>
</table>
and therefore also the energy usage to maintain a constant indoor temperature will vary over the time of a day. There are a few analytical models available to describe the dynamics of the outdoor air temperature. The most common ones are sine-exponential and sinusoidal models, based on four parameter values: sunrise \((t_{\text{sunrise}})\) and sunset \((t_{\text{sunset}})\) times and maximum \((T_{\text{max}})\) and minimum \((T_{\text{min}})\) temperature values over the day (cf. [7]). Parton and Logan [8] have shown that a sine-exponential model provides more realistic predictions, using both a sunrise and sunset time parameter value, whereas a sinusoidal model includes only the sunset time. Equation (2) presents the daytime outdoor temperature variation \(dot(t)\) and equation (3) presents the nighttime variation \(not(t)\) (adopted from [9]). The values for the time parameters are relative to midnight. Here in the evening (before midnight) \(T_{\text{min}}\) refers to the minimum temperature ahead in time and in the early morning (after midnight) it is of the day itself; similarly in the early morning (after midnight) \(T_{\text{sunset}}\) refers to the temperature at sunset of the previous day, but in the evening (before midnight) it refers to the temperature at sunset on the day itself.

\[
dot(t) = b + a \sin \left( t - 15 + \frac{t_{\text{sunset}} - t_{\text{sunrise}}}{2} \right) \left( \frac{\pi}{t_{\text{sunset}} - t_{\text{sunrise}}} \right) \tag{2}
\]

where
\[
a = \left( T_{\text{max}} - T_{\text{min}} / (1 - \sin \left( \frac{t_{\text{sunrise}} - 9}{12} \right) ) \right) \quad \text{and} \quad b = (T_{\text{max}} - a)
\]

\[
not(t) = (T_{\text{min}} - d) + (T_{\text{sunset}} - T_{\text{min}} + d e^{-\alpha(t-t_{\text{sunset})}}) \tag{3}
\]

10.2.3. Indoor temperature when cooling down takes place

When the heating system in a house is off, and the outdoor temperature is lower, the indoor temperature decreases. The rate of change of the temperature of an object is proportional to the temperature difference between it and the ambient temperature [10]. When the ambient temperature is not a constant, modeling the indoor temperature under the cooling down process is a bit challenging. To approximate the time taken to cool down overnight from a given indoor temperature to another temperature, it is crucial to know at each point in time the rate of change of the indoor temperature, thus obtaining a differential equation for \(T_{\text{id}}(t)\) which can be solved analytically. The following section adopts Newton’s law of cooling down for this; equation (4) presents the decay of indoor temperature under natural cooling down with a varying ambient temperature. Here \(k\) is the heat transfer coefficient; it consists with energy loss per degree day \(\varepsilon\) and the energy needed per degree increase of temperature (capacity \(C\)).

\[
\frac{dT_{\text{id}}(t)}{dt} = -k(T_{\text{id}}(t) - T_{\text{od}}(t)) \quad \text{where} \quad k = \frac{\varepsilon}{24C}
\]
From equation (3) the model for $T_{sd}(t)$ (focusing on cooling in the night) can be substituted:

$$\frac{dT_{id}(t)}{dt} = -k \left( T_{id}(t) - (P + Qe^{-\alpha(t-t_{sunset})}) \right)$$

where $P = (T_{min} - d)$, $Q = (T_{sunset} - T_{min} + d)$, and $t_1$ is the starting time of the cooling down process.

$$\frac{dT_{id}(t)}{dt} = kP - kT_{id}(t) + kQe^{\alpha(t_1+t_{sunset})}e^{-at}$$

$$= A + BT_{id}(t) + De^{-at}; \text{where } A = kP, B = -k, \text{and } D = kQe^{\alpha(t_1+t_{sunset})}$$

Note that when:

$$f(t) = -\frac{A}{B} + \varphi e^{Bt} - \left( \frac{D}{\alpha + B} \right) e^{-at}$$

where $\varphi$

$$= \left( f(t_1) + \frac{A}{B} + \left( \frac{D}{\alpha + B} \right) e^{-at_1} \right) e^{-Bt_1}$$

then its derivative is:

$$\frac{df(t)}{dt} = A + Bf(t) + De^{-at}$$

Therefore the differential equation for $T_{id}(t)$ can be solved analytically, thus obtaining:

$$T_{id}(t) = P + \left( f(t_1) - p + \frac{kQe^{\alpha t_{sunset}}}{\alpha - k} \right) e^{-k(t-t_1)} - \left( \frac{kQe^{\alpha t_{sunset}}}{\alpha - k} \right) e^{-\alpha(t-t_1)} \quad (4)$$

To estimate from this the time taken in the cooling down process for a given loss of temperature $\Delta T$ it is not possible to find the roots by directly solving the non-linear exponential equation in $t$ in (4). However, an adequate approach for this is to use a standard numerical method which is able to approximate roots with a sufficient accuracy. Newton’s method [11] is a good choice for this due to its simplicity and good speed (cf. [11]).

### 10.2.4. Performance of a heat pump

Heat pumps use renewable heat sources such as the ambient air of a building [12]. However, there are some drawbacks associated with heat pumps: decrease of the heating capacity and efficiency for lower ambient temperatures and accumulating frost on outdoor parts when the ambient temperature is close to 0 (cf. [13]). According to Zogou and Stamatelos [14] the energy saving gained by using a heat pump in those days was 50%. The efficiency of a heat pump is closely related to the difference between the ambient temperature (heat source) and the output temperature of the heat pump [15], in addition to other factors (cf. [14]). The performance of a heat pump is indicated by its Seasonal Performance Factor SPF in equation (5): the ratio of the heat delivered by the heat pump (energy output: $eo$) and the electrical energy supplied to it (energy input: $ei$). For air to water heat pumps $SPF$ usually
varies between 2 and 4 (e.g., for outdoor temperatures between \(-5^\circ\text{C}\) and \(15^\circ\text{C}\)) and according to Omar and Bo [13] in favorable circumstances the most efficient pumps may even show a value of 5.9. Being a dynamic property over the outdoor temperature, to approximate its value equation (6) (adopted from [9, 16]) can be used:

\[
SPF(T_{od}) = 7.5 - 0.1(T_w - T_{od})
\]

where \(T_w\) is the heating system water temperature and \(T_{od}\) is the outdoor temperature

### 10.3. The analytical model

Knowing the key factors to analyse the energy usage, the next step is to mathematically construct a model that facilitates the insight of this behavior for certain scenarios. The energy demand \((ed)\) for heating over time is an essential factor in the analysis of energy usage. It mainly concerns (1) maintaining a particular indoor temperature (thermal comfort) over a certain time period given the natural loss of heat, and (2) increasing the indoor temperature from a low value (for example, overnight) to a higher value wanted over some time period. The temperature maintenance energy demand \((tmed)\) depends on the energy loss for a given pair of indoor and outdoor temperatures (where outdoor temperature \(<\) indoor temperature): it indicates the same amount of energy through the heating system to compensate for this loss and thus maintain the given indoor temperature. The degree-days concept explained in Section 2 expresses this energy loss sufficiently [6]; the energy loss per degree-day is assumed to be \(\varepsilon\); this is different for each house/building and depends on the isolation of the border between indoor and outdoor with walls, windows, floor, roof, ventilation, etc.. For a given time interval, the value of \(tmed\) can be expressed as in equation (7). Furthermore, temperature increase energy demand \((tied)\) depends on the heat energical capacity \(C\) of the house: this indicates how much energy is needed to raise the temperature by 1 degree. Therefore \(tied\) is proportional to the temperature difference \(\Delta T_{id}\) made and relates to the notion of capacity \(C\) of the house as in the equation (8). Finally the total energy demand can be expressed as in equation (9).

\[
tmed(t_1, t_2) = \int_{t_1}^{t_2} \frac{\varepsilon}{24} (T_{id} - T_{od}(t)) \, dt \quad \text{(assuming } T_{id} > T_{od}(t))
\]

\[
tied = C \Delta T_{id}
\]

\[
ed = tmed + tied
\]

For a small time interval with length \(\Delta t\) the energy usage \(eu\) is proportional to the energy demand and relates to the seasonal performance factor \(SPF\) of the heat pump;
similarly the energy cost $ec$ is proportional to the energy usage and relates to the price $\pi_{el}(t)$ of electricity at time $t$ (see [9]) as expressed in equations (10) and (11).

$$eu(t, t + \Delta t) = \frac{ed(t, t + \Delta t)}{SPF(T_{od}(t))}$$  \hspace{1cm} (10)$$

$$ec(t, t + \Delta t) = eu(t, t + \Delta t)\pi_{el}(t)$$  \hspace{1cm} (11)$$

10.3.1. Temperature maintenance energy usage

Using equation (7) it is possible to determine an analytical expression for $tmeu(t_1, t_2)$ as follows.

$$tmeu(t_1, t_2) = \int_{t_1}^{t_2} \frac{\varepsilon}{24}(T_{ng} - T_{out}(t)) \frac{1}{SPF(T_{od}(t))} dt$$

From (3) and (6) it follows

$$tmeu(t_1, t_2) = \int_{t_1}^{t_2} \frac{\varepsilon}{24} \left( T_{ng} - \left( (T_{min} - d) + (T_{sunset} - T_{min} + d)e^{-a(t-t_{sunset})} \right) \right) \left( 7.5 - 0.1 \left( T_w - \left( (T_{min} - d) + (T_{sunset} - T_{min} + d)e^{-a(t-t_{sunset})} \right) \right) \right) dt$$

$$tmeu(t_1, t_2) = \frac{\varepsilon}{24} \int_{t_1}^{t_2} \left( a_1 + a_2 e^{-a(t-t_{sunset})} \right) \frac{1}{b_1 + b_2 e^{-a(t-t_{sunset})}} dt$$

where $a_1 = T_{ng} - T_{min} + d$, and

$a_2 = -(T_{sunset} - T_{min} + d), b_1 = 7.5 - 0.1 T_w + 0.1 (T_{min} - d), & b_2 = 0.1 (T_{sunset} - T_{min} + d) = -0.1 a_2$

$$tmeu(t_1, t_2) = \frac{\varepsilon a_1}{24} \int_{t_1}^{t_2} 1/(b_1 + b_2 e^{-a(t-t_{sunset})}) dt$$

$$+ \frac{\varepsilon a_2}{24} \int_{t_1}^{t_2} 1/(b_2 + b_1 e^{a(t-t_{sunset})}) dt$$

It is known that function $1/(a\beta) \ln(\gamma + \beta e^{at})$ is a primitive function of $1/(\beta + \gamma e^{-at})$. Therefore

$$tmeu(t_1, t_2) = \left[ \frac{\varepsilon a_1}{24} \ln(b_2 + b_1 e^{a(t-t_{sunset})}) \right]_{t_1}^{t_2}$$

$$- \left[ \frac{\varepsilon a_2}{24} \ln(b_1 + b_2 e^{-a(t-t_{sunset})}) \right]_{t_1}^{t_2}$$

From (3) it follows $e^{-a(t-t_{sunset})} = -((T_{od}(t) - T_{min} + d)/a_2)$ and therefore
10.3.2. Temperature increase energy usage

The temperature increase energy usage $tieu$ can be determined by equation (13); for the sake of simplicity it is assumed that the temperature increase will happen instantly without a time gap.

$$tieu(T_1, T_2) = \frac{C(T_2 - T_1)}{SPF(T_{od})}$$  \hspace{1cm} (13)

10.4. Using the model in heating scenarios

The analytical model can be used to analyse mathematically the energy usage over time. It represents useful knowledge to take actions to minimize the energy usage (in particular, for heating overnight as addressed here) against the ambient conditions over time. In this section the use of the model for certain heating scenarios is discussed and the results are compared to results from an alternative approach based on the simulation model presented in [16].

10.4.1. Analytical vs simulation comparison

In [16] the energy usage of an air to water heat pump (in relation to the same factors addressed in the current paper) is analysed based on simulation; therefore it is useful to compare results obtained using the analytical model introduced here to those that are obtained from the simulation model in [16]. For this purpose a scenario was chosen and analysed using both a simulation-based and an analytical approach.

The simulation model in [16] uses equations (14) and (15) for simulation through discrete time steps ($\Delta t$) of half and hour, whereas as in the analytical model introduced in this paper time is taken continuous. To improve the accuracy of the simulation model the step size was further reduced to 6mins. The behavior was analyzed specifically for data available from indoor temperature 20°C at time 21:00hrs February 1, 2012 to 06:00hrs February 2, 2012. It is assumed that the heating program is not using energy until the temperature reaches to the night goal temperature $T_{ng}$ (until then autonomous cooling down takes place). Once the indoor temperature becomes $T_{ng}$ that temperature is maintained until 06:00hrs 2nd February 2012, and the indoor temperature is increased from $T_{ng}$ to 20°C at 06:00hrs. Throughout this time interval, the outdoor temperature is assumed to be behaving as in the equation (3) and SPF is calculated as in the equation (6). According to the
collected data in [16] for this period of time, the minimum temperature $T_{min}$=-8.8°C, the temperature at sunset $T_{sunset}$=-2.52°C, the time of sunset $t_{sunset}$=17:00, and the outdoor temperature at 21:00hrs $T_{od}(21:00hrs)$=-6.6°C. Furthermore, for the remaining parameters the values were: $C=4.6$, $\varepsilon=4$, $\alpha=0.25$, $d=0.1$, $T_w=50°C$, and time step size $\Delta t = 6$ min was for the simulation.

$$T_{id}(t + \Delta t) = \max \left( T_{id}(t) - \left( \frac{\varepsilon}{24C} \right) (T_{id}(t) - T_{od}(t))\Delta t , T_{ng} \right) \quad (14)$$

$$ed(t + \Delta t) = tmed(t + \Delta t) + tied(t + \Delta t) = tmed(t) + \varepsilon(T_{id}(t) - T_{od}(t))\Delta t + tied(t) + C(T_{id}(t + \Delta t) - T_{id}(t)) \quad (15)$$

For this comparison different night goal temperatures $T_{ng}$ were selected: 14°C, 14.5°C, 15°C, 15.5°C, 16°C, 16.5°C, 17°C, 17.5°C, 18°C, 18.5°C, and 19°C. The results are shown in Fig. 1(a). According to Fig. 1(a) the analytical model gives a smooth curve (as expected), when increasing the night goal temperature (changing the night goal temperature from 15 to 19, will increase the energy usage from 22.12kWh to 24.12kWh) whereas the simulation model also provides a similar pattern but with a slight zig-zag behavior especially when the night goal temperature is high (changing the night goal temperature from 15 to 19, will increase the energy usage from 22.21kWh to 24.23kWh, so just slightly higher than for the analytical model). The main reason assumed for this difference is due to the discrete vs continuous time steps in the approaches. In simulation, in each time step, the energy usage is determined according to the $T_{id}$ and $T_{od}$ at the beginning of the time interval, and the changes during the time interval are neglected. In the simulation approach the time taken for cooling down from 20°C to $T_{ng}$ is automatically determined during the simulation, as an emergent property. Nevertheless, calculating this value is a must in the analytical approach and for that it is required to find the roots of equation (4). Due to the nature of the equation (4) it is impossible to find the roots analytically and therefore as suggested earlier Newton’s method [11] was used to obtain a close approximation of this value.

Fig. 1(b) presents the changes of temperature maintenance energy usage and temperature increase energy usage for both approaches. As per Fig. 1 (b), $tieu$ by the analytical model and by the simulation model almost has the same value for each night goal temperature (when it comes to $T_{ng} = 19°C$ the simulation data show a slightly higher value). Also $imeu$ also has somewhat equal results for both, but the simulation model shows a slight zig-zag pattern, which contributes to a noticeable zig-zag pattern in Fig. 1(a) for the total energy demand as in the equation (9)).

10.4.2. Energy usage optimization in real time

In Section 4.1 the comparison of the analytical model with a simulation model was discussed and some confidence in the results for the analytical model was
obtained. Nevertheless, in real time usage of this model it may not be possible to provide some of the parameter values. As a result of that it is important to guess/predict those values and necessary to adjust those values over time. Knowing this practical limitation, it is necessary to do the calculation of this model with certain intervals to achieve the most economical energy usage. An effective way to reduce the domestic energy usage is to let the home temperature be lower when

Fig. 1: (a) total energy usage for different night goal temperatures (b) individual energy usage (i.e. tmeu & tieu) for different night goal temperatures.
asleep. Actually, it is common among many families, to let the home temperature be a few degrees lower in the night. But, the question arises what is the appropriate temperature during the night? On the one hand, the family members are not comfortable if the home is too cold during the night, though it seems that, the lower indoor temperature during the night, the lower energy usage. A comfort range for a house is given as $T_{ng1}$ to $T_{ng2}$ (where $T_{ng1} > T_{ng2}$) and the room temperature is to be maintained within this range as in Fig. 2. From $T_{ng1}$ to $T_{ng2}$ the indoor temperature can automatically drop due to the natural cooling down. Consider a time point $t_2$ when the indoor temperature is $T_2$. At this time point there are 3 options to consider as in Fig. 2; for each option it is required to increase the temperature back to the $T_{ng1}$ at time $t_4$:

- **A**: To increase the temperature by $\Delta T$ ($< T_{ng1}$) and maintain that value until time $t_4$.
  \[ eu_{optionA} = tieu_A(t_2) + tmeu_A(t_2, t_4) + tieu_A(t_4) \]  

- **B**: To maintain temperature $T_2$ until time $t_4$.
  \[ eu_{optionB} = tmeu_B(t_2, t_4) + tieu_B(t_4) \]  

- **C**: To let the temperature $T_2$ drop until $T_{ng2}$ and then from $t_3$ to $t_4$ maintain the temperature $T_{ng2}$.
  \[ eu_{optionC} = tmeu_C(t_3, t_4) + tieu_C(t_4) \]  

As in the information in Section 2 and 3 it is possible to solve equations (16), (17), and (18) with the selected parameter values and take the option which consumes the lowest energy. This particular process can be executed in defined time intervals and in each interval it is possible to select the best option. Therefore, due to the nature of the environment, it provides the most economical solution according to the parameter values available or predicted.

### 10.5. Conclusion and future work

In this paper, the domestic energy usage by using an air-to-water heat pump was

![Fig. 2. Temperature maintains options for a night time heating program](image-url)
analysed mathematically through an analytical model. The proposed approach calculates the energy used to heat a house according a particular heating program. The analytical model was compared with a simulation model (cf. [9]). The results were almost the same, and small differences are because of time discretization in simulation. The smaller timesteps in simulation, the smaller the differences in results. The results show that there is an advantage of 7.2% in energy usage if the night temperature is left free to go down to 15 instead of 19.

In the addressed heating program, the indoor temperature is increased around the sunrise. At this time the outdoor temperature is close to its minimum. As a result, the heat pump’s performance is the lowest. So, it seems possible to save more energy by starting the heating process a little later. So, an important question which can be answered in future works is that, when is the best time for increasing the temperature, given requirements on comfort? For example, would it be feasible to postpone heating to the afternoon when the outdoor temperature (and therefore the heat pump’s efficiency) is the highest?

References


Chapter 11

Agent-Based Analysis of Annual Energy Usages for Domestic Heating based on a Heat Pump

Seyed Amin Tabatabaei, Dilhan J. Thilakarathne, Jan Treur
Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.
e-mail: s.tabatabaei@vu.nl, d.j.thilakarathne@vu.nl, j.treur@vu.nl

Abstract: This paper describes an agent-based analysis approach to determine in which way a net zero house can be obtained. In particular, it addresses agent-based simulation to estimate annual energy usage for heating based on an air to water heat pump. Based on the introduced approach house owners will be able to decide on the specifications for further renewable energy production systems to be installed, for example, solar or wind energy production systems in order to obtain a net zero house in the present and in future years.

Keywords: net zero house, annual energy usage, heat pump, SPF.

---

1 This chapter was published as:

The names of the authors are ordered alphabetically reflecting the comparable contribution of each author.
11.1. Introduction

One of the grand challenges of this time is to reduce the overall energy usage based on non-renewable sources of energy. A particular area to address this is that of domestic energy usage. In designing new houses nowadays, often the aim is to come as close as possible to an energy neutral or net zero house; e.g., [1–4]. Also for renovation of existing houses often the aim is to get close to a net zero house. A net zero house is a house that on an annual basis doesn’t use any energy. During the year there may be times that such a house still uses energy, but in the same or other periods of the year it produces energy in such a way that the net total over the year is zero. For more details and overviews of net zero houses, see [5–7]. Given the dynamics of both environmental and household factors, the emerging system is a complex system and modelling both the energy consumption and energy production over time is a research challenge.

For households in many countries much of the annual energy usage is spent on domestic heating; overall it is estimated at 43% from the total energy usage of heat related needs in the European Union in 2006 (cf. [8]). To get closer to the ideal of a net zero house, in addition to good isolation of the house, also more and more domestic heating systems are considered that allow the use of renewable energy, in contrast to most of the traditional heating systems that fully depend on non-renewable energy such as coal, gas and oil. Often heat pumps are suggested as an alternative; e.g., [9]. They take most of their energy (up to 80%) from the heat available in the ambient air, water or soil. The remaining energy usage concerns electrical energy to run the heat pump, which can be less than 40% of the energy usage based on traditional heating systems. Moreover, if the amount of electrical energy that is still needed is also produced based on renewable energy sources such as solar or wind energy in the same or different periods in a year, the total energy usage for heating can become net zero over a year. To obtain this, it is important to have an adequate estimation of the annual energy usage of a heat pump over years. Such an estimation can be taken into account when deciding, for example, for the dimensions of a solar energy production system to be installed. However, domestic heat pump users are facing a challenge to estimate their own energy usage on heating mainly because of the performance indicators given by heat pump manufacturers are far away from the dynamic conditions of using it (e.g. indoor and outdoor temperatures) and therefore, those measures are not directly helpful to plan an economical net zero house. This paper focuses on how agent-based methods can be used to get this estimation in case of an air to water heat pump (using the air as a source). Here the main agent considered is the (heat pump based) heating agent, and the main focus is on its ongoing interaction with the (strongly dynamical) environment. Other agents that play their role are energy production agents (e.g.,
based on a photovoltaic solar energy production system; e.g., [10]) and a thermostat agent.

Both the energy demand of a house and the efficiency of an air to water heat pump strongly depend on ambient temperature. This temperature varies much over the days of a year, and due to climate change it may change over the years as well. As the effect of variation of this temperature on the efficiency of a heat pump is nonlinear, simply taking average temperatures does not provide adequate estimations. In this paper, this variation of outdoor temperature is analysed and it is determined by agent-based simulation over days in different years how this variation affects the energy needed for heating. A main advantage of this approach is that by having some limited data for a short period of time the system itself is able to start to do the prediction, and over time of the system use the quality of the estimation will be naturally refined and improved. Furthermore, due to the continuous analysis of energy consumption and energy generation, this methodology includes features like monitoring the performance of the heat pump and solar panels (this is useful for prompt repairs and even for replacements), form a small scale smart grid with neighbours and further reducing the risk of higher energy demands (in special situations) with low or zero cost.

In the paper, first in Section II some background theory on heating based on a heat pump is presented. Next in Section III it is shown how parameters representing characteristics of a given house and heat pump can estimated based on empirical data on energy usage and outdoor temperature. This provides a well-tuned model of the heat pump in the given house. In Section IV based on this model the (hypothetical) energy usages for the past 10 years are analysed using empirical data on the temperature over days in these years. Section V does the same for the future 10 years, thereby using prediction models for temperature variation over days and over years. Section VI includes a discussion and future directions.

11.2. Domestic Heating by a Heat Pump

In this section, some background knowledge on domestic heating is discussed, needed to model the heating agent. The modelling approach used here is single agent-based. This agent has the responsibility to take care of the heating via the action of controlling the water temperature of the central heating system. The agent is goal-directed and reactive to information about the indoor and outdoor circumstances. This information is acquired by the agent by sensing and monitoring the outdoor temperature and the indoor and heating system water temperature. Moreover, the goal concerns the desired indoor temperature. This is obtained by communication with the human(s) in the house (via the thermostat as a communication mean). Three important elements in domestic heating are:

- The characteristics of the environment
E.g., how is the temperature over the year, how much wind is there

- The characteristics of the house
  E.g., how well isolated is the house, in how far it uses passive solar energy
- The characteristics of the heating system used
  E.g., efficiency, performance depending on circumstances

Heat pumps have often been proposed as efficient heating systems, as they use renewable heat sources such as ambient air, water or soil, and to run use only a fraction of this as electrical energy [9]. The heating capacity and efficiency of a heat pump strongly depends on the temperature of the ambient heat source used. During a frost period in a winter season this ambient temperature may become quite low, compared to milder periods [11], thus implying a lower performance in times when most heating is needed. More specifically, the efficiency of a heat pump is closely related to the difference between the ambient temperature (heat source) and the output temperature of the heat pump [11], in addition to some other factors (cf. [12, 13]). A commonly used measure on the performance of a heat pump by those manufacturers is referred as Coefficient of Performance (COP): which is the ratio of the heat delivered by the heat pump (energy output) and the electrical energy supplied to it (energy input), both measured in kWh [14]:

\[
\text{COP} = \frac{\text{energy output}}{\text{energy input}}
\]

(1)

The main concern over the COP is that it is calculated under a set of controlled conditions with defined input and output temperatures: for the European standards (EN 14511) it is to be tested at 7°C external temperature and 20°C indoor temperature, with 35°C output (hot water) from the pump [14, 15] (for American standards see [14]). These operating conditions are very different from real life situations: for example, outdoor temperature may vary widely and indoor temperature is always a subjective parameter that may include a series of values over a day. This indication is far away from a heat pump’s actual efficiency where the both indoor and outdoor temperatures are dynamically changing over the time and therefore, consumers may continuously observe that a given heat pump is consuming more electricity than they had been expected or informed. Seasonal Performance Factor SPF is a different measure for the efficiency of a heat pump which utilizes both the outdoor temperature and output temperature of heat pump into the calculation (usually this is considered for a particular period of time: for a season (e.g. winter) or a year) [14, 15]. It is the main indicator of the efficiency of a heat pump relatively with more accuracy. For air to water heat pumps in the marketplace, the Seasonal Performance Factor usually varies between 2 and 5 (e.g., for outdoor temperatures between -5°C and 15°C) [16]. Often it is around 3 (e.g., for ambient temperatures between 0°C and 10°C). Given its strong dependence on the outdoor
temperature, SPF can be approximated by a mathematical function of the outdoor temperature $T_{od}$. For this paper a linear approximation is used (adopted from [10, 17]):

$$SPF(T_{od}) = 7.5 - 0.1(T_{water} - T_{od})$$

(2)

Here $T_{water}$ is the heating system water temperature. The values of the parameters (i.e., the 7.5 and 0.1) in this approximation are in accordance with empirical data from www.liveheatpump.com (see also [10, 18]). For example, based on this function for $T_{water} = 50^\circ C$ it holds:

- SPF(10) = 3.5
- SPF(0) = 2.5
- SPF(-10) = 1.5

The energy usage also depends on the energy demand of the house. To determine the energy demand, equation (3) can be used:

$$tmed(t_1, t_2) = \int_{t_1}^{t_2} \frac{\varepsilon}{24} (T_{id}(t) - T_{od}(t)) dt$$

For 24hrs: $tmed_{avg} = \varepsilon(T_{id,avg} - T_{od,avg})$

(3)

Here $T_{od}(t)$ is the outdoor temperature at time $t$, $T_{id}(t)$ is the indoor temperature at time $t$, and $T_{id,avg}$ is the average indoor temperature over a 24 hour, and $\varepsilon$ is the energy loss per degree day (summation of individual deviations between the outdoor temperature $T_{od}$ and a given indoor temperature $T_{id,avg}$ in each time step over a 24 hours). The energy usage can be calculated from the energy demand $tmed$ and SPF by equation (4); see also [10, 18]. By averaging over the day, with $T_{od,avg}$ the average day temperature, this provides the energy usage for that day:

$$day\ energy\ usage = \frac{\varepsilon (T_{id,avg} - T_{od,avg})}{7.5 - 0.1(T_{water} - T_{od,avg})}$$

(4)

In summary, overall the following characteristics are used in this model:

- The characteristics of the environment: $T_{od,avg}$ over days
- The characteristics of the house: $T_{id,avg}$ over days, $\varepsilon$
- The characteristics of the heating system: $T_{water}$, 0.1, 7.5

All these characteristics are represented in the model (4) for the heating agent. By tuning these model parameters to a specific situation, by simulation for the 365 days of a given year (given by the 365 day temperatures), the agent’s year usage can be determined. This is what will be discussed in the next sections. Note that the first two types of characteristics depend on the house itself and on its location on the globe.
which determines its environment). The last type of characteristics is more general and independent of these, and therefore can be taken over from other situations using the same technical equipment, for example those described at www.liveheatpump.com.

11.3. Parameter Estimation for the Usage of a Heat Pump

To determine the annual energy demand of a given house over a certain time period, first it is necessary to determine the constant values of the parameters in the heating agent model expressed in equation (4) (i.e., $\epsilon$ and $T_{id_{avg}}$) with dynamic changes of $T_{od_{avg}}$. Besides, $T_{water}$ is set at 50°C. For the parameter estimation, three months (mid October 2013 to mid January 2014) of collected empirical data on daily energy usage of the heat pump in the given house (Amsterdam area Netherlands) was used together with the collected average outdoor day temperatures ($T_{od_{avg}}$) of each day for that period. In the process of parameter estimation, the sum of squares of residuals was used as the error function to be minimized. A residual is the deviation of the calculated energy usage through the model (4) on particular selected values for the above mentioned parameters, from the actual energy usage of that given day. The goal of this approach is to minimize the above mentioned error (sum of squares of residuals) with appropriate values for the parameters (least square method [19]). To implement the least square method the ‘lsqcurvefit’ function in MATLAB was used which is specifically recommended for nonlinear curve fitting [19]; a summary of the implementation is in Prog. Code 1.

It was found that the best values for the parameters are:

```matlab
% Collected data on energy usage of heat pump for domestic heating
actualEnergyUsage = [ ... ];
% Collected data of average outdoor temperature
tod_avg = [ ... ];
% To store time in days
day = 1 : 1 : length(actualEnergyUsage);
% Function F is: tmeu = (epsilon*(Tin_avg – Tod_avg)) / (7.5 - 0.1*(50 - Tod)).
F = @(x,time)(x(1).*(x(2) - tod_avg(day))) ./ (7.5 - (0.1.*((50.0 - tod_avg(day)))));
% Initial parameter values (can be any random values)
x0 = [3.5, 17];
% To find the best values for parameters which makes the sum of squares of residuals the smallest
[x, resnorm, ~, exitflag, output] = lsqcurvefit(F, x0, day, actualEnergyUsage);
```

Prog. Code 1: MATLAB implementation for parameter estimation
Fig. 1 presents two graphs of the heat pump energy usage over average outdoor temperature (part a) and over days (part b) for both the real and predicted (with the mentioned parameter values) energy usage of the heat pump. These graphs clearly show that with the parameter values as determined, the predictions mostly align with the empirical data.

The cumulative absolute error (from the residuals) is 0.2051. For all the different initial parameter value combinations that were tried, Prog. Code 1 always generated the same parameter value estimations, which gave some confidence that the obtained parameter values represent the global minimum. Thus, with these values it will be possible to predict the energy usage for the different days in any annual heating scenario for the given house with a higher degree of assurance. Though the current parameter estimation is based on three months of empirical data, this will be naturally extended in a real application mode. New data can be added at each day and the parameter estimation performed on the more extended data (for this the measurement of $T_{water}$ value also can be periodically checked and used to eliminate

![Image of two graphs](image.png)

**Fig. 1**: Energy usage of the heat pump for the given empirical data set: (a) over the average outdoor day temperature, and (b) over the days
the error from variable values of the model). Furthermore, it is clear that depending on the weather seasons (autumn, winter, spring, and summer), energy usage will be significantly different and when considering the annual energy usage it will be beneficial to find this as a combination of different seasonal energy usage values, which is confirmed through this model.

### 11.4. Prediction Based on Simulation of Past Years

The heating agent model described by equation (4) in Section 2 with the parameter values found in Section 3 enables to predict the energy usage for heating of any year for which the relevant average daily outdoor temperatures are given. As a first step, the model was used to analyse what the usages would have been for the last ten years. Moreover, it can be analysed whether in this time period a certain trend can be found in the temperature variation that may relate to climate change. If such a trend can be observed from the empirical outdoor temperature data, this will be useful for future prediction as well.

The average outdoor temperature of each day of the years 2004 to 2013 were obtained from the Royal Netherlands Meteorological Institute (KNMI) archives, in particular for the area near Amsterdam (Schiphol). Based on the daily temperature data the yearly energy usage for heating was determined by using the model described by (4). The results are shown in Fig. 2. The average year usage over these 10 years is 2674 kWh. Moreover, it was found that there is an upward trend in these annual usages, so that roughly spoken there is an increase from 2500 kWh in 2004 to 3000 kWh in 2013 (see the straight line in Fig. 2), which gives a nonneglectable difference of 20%.

To analyse the background of the observed trend over the years in usage numbers, also the daily temperatures have been inspected further. A trend in usages most probably reveals a trend in daily temperatures over the years. Indeed this turns out to be the case as is shown in Fig. 3. It was found that there is indeed a decreasing

![Fig. 2: Average outdoor temperature trend over the years 2004 to 2013](image)
trend of the yearly average day outdoor temperature from about 11.0°C in 2004 to about 10.3°C in 2013. Such a trend may indicate a (local) effect of climate change.

The additional trend information found highlights the necessity of studying the trend of temperature variation more statistically; in the next section this will be worked out for future predictions.

11.5. Prediction Based on Simulation of Future Years

In this section, it will be discussed how the domestic energy usage in the future years (2014 to 2025) can be and actually was predicted by stochastic simulation of the heating agent model and its environment. To do this, first the distribution of daily temperature in previous years is analysed, and the trend of changes was identified. In the next step, this information is used to predict daily temperatures over the future years and based on that the domestic energy usage in these years is predicted.

11.5.1. Analysing Variation in Daily Temperatures in the Last 10 years

As mentioned in Section 4, the average of the daily temperature in past 10 years has an overall decreasing trend. Upon further inspection it was found out that in the same period the variance is following an upward trend. Both effects might be local effects of climate change. Fig. 4 shows the variance of daily temperatures in the past 10 years and its upward trend.

By further analysis of the frequencies of the occurrences of the daily temperatures during last 10 years (2004 to 2013) it was found out that they can be approximated by a mixture model (cf. [20]) obtained as a weighted average of two Normal distributions \( N(\mu_1, \sigma_1^2) \) and \( N(\mu_2, \sigma_2^2) \):

\[
Mixt\text{ure Model} = 0.29\ N(16.15,8.84) + 0.71\ N(7.72,30.55)
\]
The parameters of the above formula (weight, $\mu$ and $\sigma^2$ for each normal distribution) were calculated by the Expectation Maximization (EM) algorithm described in [20]. Fig. 5 represents as squares the frequencies of all occurrences of daily temperatures in the past 10 years, obtained from empirical data. The proposed mixture model (5) is also depicted in this figure as a solid line. It should be noticed that this figure shows the diagram of 365 times the value obtained from (5), as the values of (5) are normalised at total sum 1, and here the sum is 365 days in a year.

11.5.2. Simulating Domestic Energy Usage for Future Years

After analysing the trends for average and variance of daily temperatures, and proposing a combined model for the frequency distribution of the daily temperatures in the previous 10 years, these were used to predict daily temperatures and the domestic energy usages in the future years. To this end, for each year 365 random values were generated from the model providing the distribution of daily temperatures in that year. The used temperature distribution model for each year is like formula (5) (with the same weighs), but the mean and variance of each normal distribution are changed over the years according to the mentioned trends. It is clear
that by having the daily temperature for a particular day, the energy usage of that day can be calculated from the formula (4). As a result, the energy usage of a year is estimated by summing up the estimation of energy usages in all of its days. This stochastic simulation was done 1000 times for the future years, from 2014 to 2025. Fig. 6 shows the average and standard deviation of year energy usage according to these 1000 simulations for each year.

11.6. Discussion

This paper described an agent-based approach to estimate annual energy usage for heating when using an air to water heat pump. The important feature was that the prediction will be highly sensitive to the characteristics of three basic elements: the environment, the house, and the heating agent. These all have to be taken into account rather than only the given generic Coefficient of Performance (COP) of the heat pump as provided by the manufacturer, and average outdoor temperatures. Given the motivation and trends for net zero houses, such an approach will uplift the predictions of energy demands with more confidence. Therefore, house owners will be able to decide the specifications for other renewable energy systems (e.g., solar or wind energy production systems) to obtain a net zero house. To this end, the outcomes of this approach can be combined with other energy consumption processes in the house.

More specifically, for the considered house, besides the heating per year around 1600 kWh is needed for other energy usages such as for the fridge, (dish) washing machines, light, computers, and cooking. From the outcomes in Fig. 6, it can be found that to become net zero with high probability for 2014 the total amount of compensating produced energy in 2014 should be at least 3400 kWh + 1600 kWh = 5000 kWh. For the location of the house in the Netherlands, this can be translated (based on a well-known rule of thumb) into a photovoltaic solar energy production system of 5000/0.85 = 5900 Wp (Watt peak). When, for example, solar panels are

![Estimation of Energy Usage in future years](image_url)
used of 250 Wp, this implies that 24 of such panels are needed. However, when not only for 2014, but also for the years up to 2024 a net zero situation is aimed for, a bit more is needed (assuming that the 1600 kWh will not be reduced over time). In that case, the expected maximal overall energy usage per year is $4000 \text{ kWh} + 1600 \text{ kWh} = 5600 \text{ kWh}$. This can be translated into $\frac{5600}{0.85} = 6600 \text{ Wp}$ needed, which can be obtained by 27 solar panels (of 100 cm x 165 cm) of 250 Wp. A more refined calculation might also take into account on the one hand that solar panels may become slightly less efficient over the years (e.g., by 5 or 10%), but on the other hand, also more advanced types of panels may be considered, for example of 270 Wp.

In a further practical setup based on this approach, three agents and their dynamical environment can be considered: a heat pump based heating agent, a thermostat agent, an energy production agent, and their dynamical environment with its variation in outdoor temperatures and other elements. The heating agent should generate energy for heating each day, whereas the thermostat agent monitors indoor and outdoor temperature of these days in parallel thereby using necessary sensory equipments and access to data sources related to the location of the house. By enabling a communication among these agents and interaction with the environment, the approach described in this paper will enable them to determine parameter values representing characteristics that are particular to that domestic heating energy usage behaviour. Furthermore, the thermostat agent will be able to analyse the trends of the outdoor temperatures by access to the relevant data sources specific to that location of the house. As shown in this paper, these agents together will be able to predict the annual energy usage for heating specific to that context, and based on this provide advice to the house owner.

The approach discussed in the previous sections is agent based approach based on an agent specification by mathematical equations. Alternatively, the agent specification could be expressed in a rule-based format as is more common. However, semantically it would be the same agent. One of the reasons for the agent paradigm is the future extension of this work that will include energy usage optimization by allowing a communication between these agents and also with other agents which are necessary with entities essential in energy management (including the behaviour aspects of the human agents involved). Energy related systems are naturally complex systems where the logical stability of the system is always far away from equilibrium and it is necessary to engage from situation to situation with enough details of data through a continuous monitoring, communication and tuning process for a rational optimisation. For example, when considering the temperature settings of a house from evening to morning (for the winter season) for a comfortable stay and sleep, it is a question at what time the heat pump should start and how to control the temperature values over time so that the most economical energy usage
with a satisfactory level of comfort is obtained (see [21]). This types of questions can be answered through agent based approaches more easily and proactively. Having a smart household agent based energy management system will not only provide more realistic predictions for annual energy usage but also:

- enable to construct detailed profiles of energy usage of a given household (which includes sufficient information to identify patterns of energy usage that may improve with various adaptation techniques such that energy usage will be minimized)
- promptly identifies problems of devices (e.g. heat pump, solar panel, etc.) and to do the necessary repairs or replacements to fulfil the needs for a net zero house
- extending this system with systems describing neighbouring houses in order to build a coherent system that includes mutual benefits (for example due to sudden energy demand needs it maybe possible to get the energy from others with low or zero cost, and further to move forward for a small scale wind plant that can be shared).

References


Chapter 12

Methods for a Smart Thermostat to Estimate the Characteristics of a House Based on Sensor Data\(^1\)

Wim van der Ham\(^{a}\), Michel Klein\(^{b}\), Seyed Amin Tabatabaei\(^{b}\), Dilhan J. Thilakarathne\(^{b}\), Jan Treur\(^{b}\)

\(^{a}\)Quby, Joan Muyskenweg 22, 1096 CJ Amsterdam, The Netherlands

\(^{b}\)Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.
e-mail: s.tabatabaei@vu.nl, d.j.thilakarathne@vu.nl, j.treur@vu.nl

Abstract: Smart thermostats can play an important role in achieving more economic energy usage in domestic situations. This paper focuses on the energy used for natural gas-based heating, and monitoring of gas usages versus indoor and outdoor temperatures over time. Two methods are presented that enable the smart thermostat to learn, over time, characteristics of the house such as heat loss rate and heat capacity. Through this, the thermostat can make some homeowners aware, for example, that there is room for improvement in insulation of the house. The presented methods are able to deal with sensor data with varying extents of imperfection concerning their completeness.

Keywords: heat loss rate; heat capacity; cooling down rate; heating degree days, smart thermostat.

\(^1\) This chapter was presented at the International Conference on Environmental and Climate Technologies and will appear in:
The names of the authors are ordered alphabetically reflecting the comparable contribution of each author.
12.1. Introduction

In northern parts of Europe much energy is used for domestic heating. Smart thermostats are devices used more and more to get insight in domestic energy usage; e.g., [1]–[3]. They can contribute to more economic energy use in two ways. On the one hand, they can control the heating system in an economic manner, for example by not heating the house if they detect that nobody is at home. But, on the other hand, they play an important role in making the persons in the house aware of their energy usage; so, that they are encouraged to change their behaviour or their house into a more sustainable situation.

A challenge here is to give the smart thermostat the intelligence needed to analyze what are the energetic characteristics of the house. This could provide the basis for comparisons with other users and tailored advices about measures to reduce energy usage. To achieve this, both monitoring devices and analysis methods for interpreting the data are used. The work described in this paper uses the data of gas consumption used for heating and the indoor and outdoor temperatures over time to derive the energetic characteristics of the house. A particular additional challenge is that often sensor data do not provide complete time series. From time to time data may be missing. To analyse the heating characteristics of a house, often so called heating degree day methods are used; e.g., [4]–[9]. These methods allow to estimate how the heating demand of a house relate to differences between indoor and outdoor temperatures.

Taking such heating degree day and other information, in this paper two methods are discussed that enable the smart thermostat to estimate characteristics of the house such as heat loss rate and heat capacity; so, that home owners can possibly be made aware that there is room for improvement in insulation of the house. The proposed methods are also able to handle sensor data with certain extents of imperfection concerning their completeness.

In the paper, in Section 2 some background theory about gas-based heating is discussed. In Section 3, the available dataset is described. Section 4 describes how the adaptive methods have been applied to the dataset and which outcomes were achieved. Finally, Section 5 is a discussion.

12.2. Theoretical basis

This section discusses the theoretical background concerning gas-based heating and the relation to outdoor temperature, and central concepts characterizing a house such as heat loss rate $\varepsilon$ (depending on insulation level of the house) and heat capacity $C$ of the house (depending on volume of the house).
12.2.1. Central concepts

The energy loss of a house per time unit with indoor temperature $T_{id}$ and lower outdoor temperature $T_{od}$ is proportional to the temperature difference between indoor and outdoor temperatures. The proportion factor $\varepsilon$ is the loss rate:

$$\text{energy loss per time unit} = \varepsilon(T_{id} - T_{od}) \quad (1)$$

This loss rate depends on insulation level of the area of the house in contact with the outside: the walls, windows, floor and roof.

Another central concept is the heat capacity of a house. The amount of energy needed to increase the indoor temperature is proportional with the difference $\Delta T_{id}$ in indoor temperature. The proportion factor $C$ is the heat capacity:

$$\text{energy needed for increase} = C\Delta T_{id} \quad (2)$$

The heating capacity of a house, $C$, depends on the volume (content) of the house. Note that, during a time interval of temperature increase, still energy loss takes place as well. More specifically, there are three types of situations: A) Indoor temperature increase: In periods of increase of indoor temperature both types of energy described by (1) and (2) have to be added to each other to get the total amount of energy spent. B) Maintaining constant indoor temperature: When heating takes place just to maintain a given indoor temperature, only the energy loss is compensated by the heating. The amount of energy needed for this is described by (1). C) Natural cooling down without heating: When no heating takes place the house follows a natural cooling down process. During such a time interval the energy loss (per time unit) described by (1) leads to a temperature decrease (per time unit) described by (2):

$$\varepsilon(T_{id} - T_{od}) = \frac{C\Delta T}{\Delta t} \quad \text{Or} \quad \frac{\Delta T_{id}}{\Delta t} = \frac{\varepsilon}{C}(T_{id} - T_{od}) \quad (3)$$

So in a cooling down process, the speed of indoor temperature decrease is proportional to the difference between indoor and outdoor temperature. This proportion factor $\varepsilon/C$ is called the cooling down rate, indicated by $\mu$.

The concepts discussed above can be applied for specific time instants or very short time durations, but they can also be used for longer time periods, such as hours, days, months, seasons or years. In the latter case, the formulae (1), (2) and (3) can still be applied but some form of summation or integration over time is needed. This has been done in the two approaches that have been developed. The two proposed approaches put different requirements on the data to which they can be applied.
12.2.2. First Approach

The first approach is based on a calculus of what is called (heating) degree days. Degree days based energy consumption analysis is a well-known approach to quantify the relation between energy usage and outdoor and indoor temperatures; e.g., [4]–[9]. Through this, it is possible to approximate energy consumption and energy performance of a building based on historical data and that can be used to estimate the energy loss per degree per day (kWh/°C day). In this approach, only the energy which needed to compensate for the energy loss of the house is counted. Therefore (1) above applies, but still an integration process over a longer time period has to be applied to it. Since both indoor and outdoor temperatures are varying, a specific adapted definition of a number of degree days (dd) during a 24-hour period can be expressed mathematically in the following manner:

\[ dd = \int_{0}^{24 \text{ hrs}} (T_{in}(t) - T_{od}(t)) \, dt, \quad \text{when } T_{in}(t) > T_{od}(t) \text{ for all } t \] (4)

Here, \( T_{in}(t) \) and \( T_{od}(t) \) are indoor and outdoor temperatures as function of time \( t \). In this approach, the quality of calculating degree-day value depends on the quality of indoor and outdoor air temperature values. In practical applications, this equation can be transformed into a more discrete version as:

\[ dd = \sum_{\Delta t \text{ intervals}} (T_{in}(t) - T_{od}(t)) \] (5)

The smaller the \( \Delta t \) values, the higher the accuracy (in our experiments, \( \Delta t \) was taken as one hour). When energy usage data for heating are available, based on (1) above the relation between heating energy usage and the number of degree days can be mathematically expressed as:

\[ \text{energy usage} = \varepsilon dd \] (6)

Here, \( \varepsilon \) is the energy loss rate that gives a good indication about quality of insulation of a house. When \( \varepsilon \) has a low value, that house has good insulation, but high values indicate bad insulation. This can be used as an indication to take initiative to enhance the awareness of householders to take necessary actions.

12.2.3. Second Approach

In the second approach, the focus is not (only) on the energy loss over 24 hour periods, but also more in particular on analysing the shorter time periods in which the indoor temperature has a downward or upward trend. For example, in the morning many houses are heated from a lower night indoor temperature to the higher day indoor temperature. This approach enables to estimate not only the loss
rate $\varepsilon$ but also the capacity $C$, and the cooling down rate $\mu$. For cooling down periods, when no heating energy is given to the house, by integration of (3) the equation (7) and (8) can be obtained. Furthermore by combining (7) and (8) we get (9).

\[ \text{energy loss} = C \Delta T_{in} \]  \hspace{1cm} (7)

\[ \text{energy loss} = \varepsilon \int (T_{in} - T_{od}) dt \]  \hspace{1cm} (8)

\[ \frac{\varepsilon}{C} = \frac{\Delta T_{in}}{\sum(T_{in} - T_{od})\Delta t} \]  \hspace{1cm} (9)

As mentioned, $\varepsilon/C$ is the cooling ratio $\mu$, and shows the rate or speed of decreasing the indoor temperature of a house when no heating energy is provided. On the other hand, for the periods of time in which heating energy is provided to increase the indoor temperature, by integration from (1) and (2) the following is found:

\[ \text{energy demand} = \text{energy for maintaining a temperature} \]
\[ + \text{enrgy for increasing temperature} \]
\[ = \varepsilon \int (T_{in} - T_{od}) dt + C \Delta T_{in} \]  \hspace{1cm} (10)

In the second approach, equations (9) and (10) are used to calculate the capacity $C$, cooling down rate $\mu$ and loss rate $\varepsilon$ for a house. To do this, four main steps are used: 1) Data extraction to find the time intervals in which upward or downward trends in indoor temperature occur. 2) Determining cooling down rate $\mu$ for the time intervals in which the house is cooling down (based on equation (9)). This also provides one relation $\varepsilon = \mu C$ between the capacity $C$ and loss rate $\varepsilon$. 3) Determining another relation between the capacity $C$ and loss rate $\varepsilon$ using the time intervals in which the indoor temperature (monotonically) increases (based on equation (10)). 4) Determining the values for the capacity $C$ and loss rate $\varepsilon$ by combining the two relations obtained in 2 and 3. Details of each step are described below:

**Data Extraction and Cleaning:** Since this article is on the space heating usage, the focus was on the colder season, which for the available dataset was limited to most of the days in March and some days in April. As a first step, data of hours which have all required information were extracted. The required information for an hour are: Gas usage in that hour, Indoor temperature per minute, Outdoor temperature for three time points (at minutes 5, 25 and 55). For the outdoor temperature in other minutes of an hour, a linear interpolation was used.

**Calculating the cooling down rate, $\mu$:** To do this, for each house the set of continuous hours with no gas usage and a downward trend for the indoor temperature are extracted. As mentioned, to calculate the cooling down rate, it is
necessary to analyse time intervals in which no heating energy is provided by heating system. To get rid of delays in the heating and metering system (it takes some time after gas usage to deliver the energy to the house environment through the hot water inside the radiators and the meter only provides data per hour), the first hour of each set of continuous hours is removed, and the cooling down rate $\mu$ is determined according to equation (9). Given the number of time intervals in which cooling down takes place, for each house several estimated values for the cooling down rate are calculated, some of which are outliers. When, for example, windows or doors are open, the energy loss will increase significantly. In this condition, (9) does not apply anymore, and the value of the calculated rate will be much higher than the real one. However, it can still be used as a technique in smart thermostats, to alert the residence that probably a window or door is remained open for hours. So, to remove the effect of such outliers on the result, the median (not average) of these values as the cooling down rate of the house is used.

**Calculating loss rate $\varepsilon$ and capacity $C$:** To do this, the time intervals are addressed in which the heating system is activated (gas usage > small threshold) and the indoor temperature has an upward trend. As the capacity has an effect only when the temperature is changing (not when it is constant), the focus is on the time intervals in which the indoor temperature is increasing (from when indoor temperature reaches its local minimum, till 2 hours after that). The amount of energy usage is calculated from (10). So, by combining the outcomes of (10) and the cooling down rate $\mu$, the following is obtained:

$$C = \frac{\text{energy usage}}{\mu \int (T_{in} - T_{od}) dt + \Delta T_{in}}$$  \hspace{1cm} (11)

On average, the efficiency of natural gas based heating systems in North Western Europe is taken as 7 kWh/m³; for example, see [http://nl.wikipedia.org/wiki/Aardgas](http://nl.wikipedia.org/wiki/Aardgas). Using this, the energy usage based on the amount of gas used is calculated by (12). Again, the median of the obtained capacities and of $\varepsilon$ is chosen.

$$C = \frac{7 \times \text{gas usage}}{\mu \sum (T_{in} - T_{od}) \Delta t + \Delta T_{in}}$$  \hspace{1cm} (12)

**12.3. Dataset**

The data that was used to validate the above methods was collected by a smart thermostat company in the period starting from the 26th of March 2014 until the 23th of July 2014. During this period, data was collected from 67 households that have a smart thermostat (Toon) installed at home and agreed to be part of this data collection. Depending on the boiler type and the software installed, the number of
collected variables is between 10 and 35. For this study, only the following
variables are used that were collected from the thermostat: Set point:- value of the
goal temperature set in the thermostat, Indoor temperature:- actual indoor
temperature as measured by the thermostat; Gas use:- total gas used. The set point
and indoor temperature have a value for every minute in degree Celsius. Gas use is
in liters (0.001 m³) and has a value for every hour. The outdoor temperature is
missing from this data because the thermostat does not directly measure this value.
These temperatures were acquired using the Weather Underground API and the
postal code of the user to retrieve the location. The dataset was used in an
anonymized way. The data that was collected from the thermostat was combined
with general information that is known about the characteristics of the household.
These include the construction period of the house divided in the following six
groups: before 1946, between 1946 and 1964, between 1965 and 1974, between
1975 and 1987, between 1988 and 1999, after 1999. The total floor area in square
meters of the house is represented in the following eight groups: between 50 and 75,
between 75 and 100, between 100 and 125, between 125 and 150, between 150 and
200, between 200 and 300, between 300 and 500, more than 500, all in square
meters. The type of household is in the dataset used for this research one of the
following: apartment, terraced, semidetached or detached. And, the last
characteristic of the household is the number of people that live in the house, which
can be: one to four, or more than four.

12.4. Experimental analysis

In this section, we describe how the methods introduced in Section 2.2 and 2.3
have been applied to the data set. Approach 1, based on the degree-days, depends on
the integration of consecutive data points for the indoor and outdoor temperature.
Approach 2 requires the identification of specific intervals in which heating takes
place and intervals in which no heating takes place.

The main research question in this paper is whether the heating characteristics of
a house can be estimated based on temperature and gas usage data, given an
imperfect and incomplete data set. A number of hypotheses were defined:

H1. The calculated capacity C of a house is positively correlated with the size
of a house (a large house requires more energy to be heated).

H2. The calculated loss rate ε is correlated with insulation level of the house.
As the data set does not contain specific information about this, we
operationalize this hypothesis in the following ways:

H2a. The calculated loss rate ε is correlated with the building year of the
house: the older the house, the higher the loss rate.

H2b. The calculated loss rate ε is correlated with the type of the house: the
more detached a house, the higher the loss rate.

H3. The calculations of the loss rate ε via each of the approaches will result in
similar values.
Approach 2 will be better applicable to incomplete data, as it requires smaller intervals of consecutive data.

It turned out that the dataset of indoor temperature and gas usage of 67 households had a lot of missing values. For approach 1, only 40 days were identified in which there were sufficient consecutive data points to calculate a degree day value. These 40 days belonged to 12 households, resulting in 2 till 4 usable days for each of the 12 households. Table 1 shows the resulting loss rate calculations for Approach 1.

Applying approach 2 resulted in 166 usable intervals belonging to 11 different households. The number of respective cooling and heating intervals are shown in Table 2. The outcomes of approach 2 are listed in Table 3.

<table>
<thead>
<tr>
<th>House#</th>
<th>loss rate ε</th>
<th>Building year</th>
<th>Type</th>
<th>Size category (m2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.83</td>
<td>1988_2000</td>
<td>Semi Detached (SD)</td>
<td>125_150</td>
</tr>
<tr>
<td>2</td>
<td>2.92</td>
<td>1975_1988</td>
<td>Semi Detached (SD)</td>
<td>125_150</td>
</tr>
<tr>
<td>3</td>
<td>4.68</td>
<td>1946_1965</td>
<td>Detached(D)</td>
<td>200_300</td>
</tr>
<tr>
<td>4</td>
<td>3.89</td>
<td>After 2010</td>
<td>Terraced (T)</td>
<td>150_200</td>
</tr>
<tr>
<td>5</td>
<td>3.58</td>
<td>1946_1965</td>
<td>Terraced (T)</td>
<td>125_150</td>
</tr>
<tr>
<td>6</td>
<td>4.98</td>
<td>Before 1946</td>
<td>Terraced (T)</td>
<td>200_300</td>
</tr>
<tr>
<td>7</td>
<td>10.47</td>
<td>After 2010</td>
<td>Terraced (T)</td>
<td>150_200</td>
</tr>
<tr>
<td>8</td>
<td>3.36</td>
<td>Before 1946</td>
<td>Apartment (A)</td>
<td>075_100</td>
</tr>
<tr>
<td>9</td>
<td>2.92</td>
<td>1965_1975</td>
<td>Terraced (T)</td>
<td>125_150</td>
</tr>
<tr>
<td>10</td>
<td>5.02</td>
<td>1946_1965</td>
<td>Terraced (T)</td>
<td>150_200</td>
</tr>
<tr>
<td>11</td>
<td>3.00</td>
<td>1975_1988</td>
<td>Semi Detached (SD)</td>
<td>125_150</td>
</tr>
<tr>
<td>12</td>
<td>3.63</td>
<td>1988_2000</td>
<td>Terraced (T)</td>
<td>150_200</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>House#</th>
<th>cooling intervals</th>
<th>heating intervals</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>71</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>5.92</td>
<td>7.92</td>
<td></td>
</tr>
</tbody>
</table>

H4. Approach 2 will be better applicable to incomplete data, as it requires smaller intervals of consecutive data.
12.4.1. Relation between capacity and size of the house

The size of the houses in the dataset is represented by four categories. Only approach 2 results in estimate of the capacity. Fig. 1 shows the correlation between the capacity and the size of the households.

There is only one household in the smallest category. This one house has a very high capacity. The mean value of the capacity for the category A100_150 (5 items) is 4.3, while for A150_200 (4 items) the value is 5.6. The average value of the capacity of the houses in the largest category A200_300 is 7.2.

12.4.2. Relation between loss rate and type of the house

Both approach 1 and approach 2 can be used to determine the relation between the loss rate $\varepsilon$ and the type of house. Fig. 2 shows the derived loss rates for the different types of houses. It can be seen that the average loss rate calculated via approach 1 is higher than the one calculated by approach 2. Also, it is visible that the ranking is different. Calculating the averages of the loss rates per house type leads to the following results (Table 4). The only instance of the type “apartment” scores relatively high. The average of the “terraced” is slightly lower than the average of “semi_detached”. This is in line with the fact that a terraced house has a higher percentage of walls with contact with the outside.

Another aspect of the type of house is the building year. The same analysis has been performed as for the building type. Fig. 3 shows the relation for both approaches. There is no clear correlation visible for approach 1. For approach 2, the

<table>
<thead>
<tr>
<th>Type</th>
<th>Average $\varepsilon$</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apartment</td>
<td>3.6</td>
<td>1</td>
</tr>
<tr>
<td>Terraced</td>
<td>1.8</td>
<td>3</td>
</tr>
<tr>
<td>Semi Detached</td>
<td>2.0</td>
<td>7</td>
</tr>
<tr>
<td>Detached</td>
<td>4.0</td>
<td>1</td>
</tr>
</tbody>
</table>
newer buildings seem to have lower loss rates. When the average loss rates per category are calculated, this becomes even clearer. Fig. 4 depicts the average loss rates per category. Except for one category (based on one house), it holds that newer buildings have lower loss rates.

With respect to the hypotheses, the following conclusions can be drawn. It appears that the outcomes of approach 1 are less consistent with the expected results than the outcomes of approach 2. Also, the number of useful segments of data is higher for approach 2 than approach 1. Both facts support H4: approach 2 is better for handling incomplete data. Related to this, it has been shown that the outcomes of the calculations of the loss rates were quite different for approach 1 and 2. H3 therefore has to be rejected: the outcomes are not similar. Hypothesis H1 and H2 are largely validated for approach 2. The average values of the capacity per size category on the one hand and the loss rate per type of house and per building year on the other hand increase for each higher category, except for one single measurement in each analysis. It has to be noted, however, that the data set is too small to draw strong conclusions.

12.5. Discussion

In this paper, it has been addressed how a smart thermostat (cf. [1]–[3]) can be equipped with the means to estimate in an adaptive manner some of the characteristics of a house over time: heat loss rate, heat capacity and cooling down rate. Two approaches have been presented one of which is based on heating degree day analysis methods; cf. [4]–[9]. The approaches have been tested using a data set including data of a variety of houses of different types. This dataset shows a certain extent of incompleteness. The approaches both turn out to be able to handle data with some extent of incompleteness, where the second approach seems to do this a bit better than the first one. The proposed methods actually can be incorporated in a smart thermostat to make it adaptive over time, and smarter due to that.

There are few commercial products (e.g., Joulo and Nest) available mainly to

![Fig.3: Relation between loss rate and type of house (above: approach 1, below approach 2)]
provide personalised home heating advice to households. Joulo [10] has used a model-based approach for this purpose, which does not use gas or electricity consumption of the heating system in the calculations but only indoor and outdoor temperatures. However, this technique is based on some unrealistic assumptions. The first assumption is that "the thermostat has a single set-point through a day", which in most of the real houses is not the case. Another (and more important) unrealistic assumption is that “energy produced by heating system in all intervals is always the same”. Even if the heating system has just two modes (ON/Off) this cannot be a true assumption, since the system may be ON just for part of an interval.

In the proposed techniques, focus is to calculate the characteristics of a house based on used heating energy and to improve the awareness of users about impact of the current state of their house on heating energy cost. From an application perspective, these may be useful features on specific thermostats and maybe it is useful to integrate features of most of these commercial products to improve the sustainability of energy usage. Another interesting feature is to also incorporate dynamic pricing (especially with smart grid based systems) in the calculations; cf. [11].

References


Chapter 13

Cognitive Simulation Driven Domestic Heating Energy Management

Dilhan J. Thilakarathne, Jan Treur

Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.
E-mail: d.j.thilakarathne@vu.nl, j.treur@vu.nl

Abstract: Energy management for domestic heating is a non-trivial research challenge, especially given the dynamics associated to indoor and outdoor air temperatures, required comfortable temperature set points over time, parameters of the heating source and system, and energy loss rate and capacity of a house. In addition to all these factors, human influence or interaction is also a key aspect in this complex system. It is difficult and very costly to conduct experiments of this nature to scrutinize the dynamics and optimal efficiency of the system under all circumstances and constraints. This paper focuses on a domestic heating energy management system using an air to water heat pump and uses a pre-developed mathematical model for its performance. This mathematical model is integrated within a computational dynamic cognitive model which was developed based on neurocognitive evidence. An approach like this can be used as an experimental workbench for complex scenarios.

Keywords: Heating energy management; Cognitive modelling; Mathematical model; Heat pump; Short term and long term goals.

---

1 This chapter was presented at the 28th International Conference on Improving Sustainability Concept in Developing Countries and will appear in: Thilakarathne, D., & Treur, J. (to appear). Cognitive Simulation Driven Domestic Heating Energy Management. In Proceedings of the 28th International Conference on Improving Sustainability Concept in Developing Countries. Elsevier. The names of the authors are ordered alphabetically reflecting the comparable contribution of each author.
13.1. Introduction

Energy is an important ingredient for the development of a country and adequate energy management is a vital aspect for a sustainable development. In many countries most of the energy usage goes to domestic heating and cooling [1, 2]. The amount of energy that can be saved by properly managing the energy usage of a house may not be felt as very significant, but in a global context it surely may provide a significant contribution to a better environment in the future. It has been observed that a major portion of energy used for domestic heating and cooling is needlessly wasted [2]. Therefore, many researchers focus on how to reduce domestic energy usage, specifically for heating related energy usage. The research reported here addresses different aspects of energy saving, among others making use of predicting variations in outside temperature, characteristics of a house, user preferences, improved devices, and usage of renewable energy sources. This paper mainly focuses on domestic heating management through an air to water heat pump together with relevant human cognition. For performance of an air to water heat pump a mathematical model was already developed [3]. This model describes the performance of an air to water heat pump, over time, in relation to variation both for indoor and outdoor temperatures and characteristics of a house. This model has been validated with realistic simulations and partly with empirical data [3,4].

Due to the developments in brain imaging and recording techniques, the insight in human brain processes is growing rapidly. This contributes to an improved quality of relevant data and knowledge, and to the development of new methods to explore this most complex system within human anatomy through different dimensions. Based on findings from Cognitive Neuroscience, a dynamic computational model was developed for action selection taking into account situation awareness [5]. This model includes a number of relevant cognitive states: performative desires, subjective desires, feeling, action preparation, ownership, attention, intention, and awareness. It models the interplay between conscious and unconscious processes. Behaviour of inhabitants in a house has a strong impact on the energy consumption and it is an important factor for energy waste reduction, especially in a dynamic context [6].

This paper combines the mathematical model that was developed for analysis of smart daily energy management for an air to water heat pump with the cognitive model for situation awareness. It is a common problem in energy related research that practical experiments cost a lot in effort and money. Furthermore, when it is required to integrate human cognition also into the experiment it is much harder. Having this type of research contributes to facilitate a sophisticated working environment to simulate complex situations including behaviour of both technical systems and humans. This paper focuses on a single house with a human’s desires and intentions for selecting night indoor temperature. An inhabitant has desires on a
comfort level required at night, strong desires to save money, and has to balance between comfort level and money (or energy) saving.

13.2. A mathematical model for heating by a heat pump and a cognitive model for action selection

This paper integrates two models that have been developed earlier. The first explains a mathematical model for the domestic energy usage by using an air-to-water heat pump. Due to space limitations only a brief description of this model will be presented in the current paper; more details can be found in [3]. The model calculates the energy used to heat a house over time according to a particular heating program. The second model covers cognitive processes behind action selection taking into account situation awareness; it has been applied to analyse the role of situation awareness in the aviation domain in [5]. This model addresses the interplay of a bottom-up and a top-down process and in particular how they interact with each other in order to cognitively drive a given situation to an adequate action.

13.2.1. Mathematical model for the domestic energy usage by an air-to-water heat pump

For a mathematical model of domestic energy usage by an air to water heat pump there are various dynamic factors to be considered. Among them one is the dynamics of outdoor air temperature. This also further can be divided into two parts daytime and night-time outdoor air temperature. The outdoor air temperature typically shows a 24h periodic behaviour (e.g., [7, 8]) and therefore also the energy usage to maintain a constant indoor temperature will vary over the time of a day. Sine-exponential and sinusoidal models are the most common analytical models used for this purpose. They calculate the pattern of outdoor temperatures over a day (and night) based on four parameters: sunrise time ($t_{\text{sunrise}}$) and sunset time ($t_{\text{sunset}}$), and maximum ($T_{\text{max}}$) and minimum ($T_{\text{min}}$) temperature values over the day [7]. Note that these are parameters for one given day, but they themselves are in fact variables at a time scale over a year, as they differ from day to day.

Equation (1) below presents the daytime outdoor temperature variation $dot(t)$ and equation (2) presents the night-time variation $not(t)$. The values for the time parameters and variables are relative to midnight. Here in the evening (before midnight) $T_{\text{min}}$ refers to the minimum temperature ahead in time and in the early morning (after midnight) it is of the day itself; similarly in the early morning (after midnight) $T_{\text{sunset}}$ refers to the temperature at sunset of the previous day, but in the evening (before midnight) it refers to the temperature at sunset on the day itself.

\[
dot{t} = b + a \sin \left( \left( t - 15 + \frac{t_{\text{sunset}} - t_{\text{sunrise}}}{2} \right) \frac{\pi}{t_{\text{sunset}} - t_{\text{sunrise}}} \right) \quad (1)
\]
where \( a = \left( T_{\text{max}} - T_{\text{min}} / (1 - \sin \left( \left( t_{\text{sunrise}} - 9 \right) \frac{\pi}{12} \right) \right) \) and \( b = (T_{\text{max}} - a) \)

\[
\not(t) = (T_{\text{min}} - d) + (T_{\text{sunset}} - T_{\text{min}} + d) e^{-a(t-t_{\text{sunset}})}
\] (2)

Another important element for the heating model is the notion of degree days \((dd)\). This concept has been introduced to approximate the analysis of energy consumption and energy performance of a building based on historical data (e.g., [9]). The number of degree-days is defined as the summation of individual deviations between the outdoor temperature \(T_{\text{od}}\) and a given indoor temperature \(T_{\text{id}}\) in each time step for a time interval from \(t_1\) to \(t_2\). This can be expressed mathematically as:

\[
dd(t_1, t_2) = \int_{t_1}^{t_2} (T_{\text{id}} - T_{\text{od}}(t))dt \quad \text{when } T_{\text{id}} > T_{\text{od}}(t) \text{ for all } t \text{ with } t_1 \leq t \leq t_2
\] (3)

The next important process is natural indoor cooling down process when the heating system is off (where outdoor temperature is lower). The rate of change of the temperature of an object is proportional to the temperature difference between it and the ambient temperature [10]. When the ambient temperature is not a constant, modelling the indoor temperature under a cooling down process is a bit challenging. To approximate the time taken to cool down overnight from a given indoor temperature to another temperature early in the morning, it is crucial to know at each point in time the rate of change of the indoor temperature, thus obtaining a differential equation for \(T_{\text{id}}(t)\) which can be solved analytically or numerically. The following section adopts Newton’s law of cooling down for this; equation (4) presents the decay of indoor temperature under natural cooling down with varying ambient temperature. Here \(k\) is the heat transfer coefficient; it consists with energy loss per degree day \(\varepsilon\) and the energy needed per degree increase of temperature (capacity \(C\)). The detail steps of deriving the equation (4) can be found in [3].

\[
T_{\text{id}}(t) = P + \left( f(t_1) - P + \frac{kQe^{\alpha t_{\text{sunset}}}}{\alpha - k} \right) e^{-k(t-t_1)} - \left( \frac{kQe^{\alpha t_{\text{sunset}}}}{\alpha - k} \right) e^{-a(t-t_1)}
\] (4)

where \(k = \frac{\varepsilon}{2AC}\), \(P = (T_{\text{min}} - d)\), \(Q = (T_{\text{sunset}} - T_{\text{min}} + d)\),
\(t_1\) is the starting time of the cooling down process,
\(f(t) = -\frac{A}{B} + \varphi e^{\beta t} - \left( \frac{D}{\alpha + B} \right) e^{-\alpha t},\)
\(\varphi = \left( f(t_1) + \frac{A}{B} + \left( \frac{D}{\alpha + B} \right) e^{-\alpha t_1} \right) e^{-Bt_1},\)
\(A = kP, B = -k, \text{ and } D = kQe^{\alpha(t_1+t_{\text{sunset}})}\)
To estimate from this the time taken in the cooling down process for a given loss of temperature $\Delta T$ it is not possible to find the roots by directly solving the non-linear exponential equation in $t$ in (4). However, an adequate approach for this is to use a standard numerical method which is able to approximate roots with a sufficient accuracy. Newton’s method [11] is a good choice for this due to its simplicity and good speed.

Finally it is important to understand the performance of the heat pump. The performance of a heat pump is indicated by its Seasonal Performance Factor $SPF$ in equation (5): the ratio of the heat delivered by the heat pump (energy output: $eo$) and the electrical energy supplied to it (energy input: $ei$). Being a dynamic property over the outdoor temperature, to approximate its value in a linear manner, equation (6) (adopted from [9, 16]) can be used.

$$SPF = \frac{\text{energy output}}{\text{energy input}} = \frac{eo}{ei}$$  \hspace{1cm} (5)

$$SPF(T_{od}) = 7.5 - 0.1(T_w - T_{od})$$  \hspace{1cm} (6)

where $T_w$ is the heating system water temperature and $T_{od}$ is the outdoor temperature.

The energy demand ($ed$) for heating over time is an essential factor in the analysis of energy usage. It mainly concerns (1) maintaining a particular indoor temperature (thermal comfort) over a certain time period given the natural loss of heat, and (2) increasing the indoor temperature from a low value (for example, overnight) to a higher value wanted over some time period. The temperature maintenance energy demand ($tmed$) depends on the energy loss for a given pair of indoor and outdoor temperatures (where outdoor temperature < indoor temperature): it indicates the amount of energy through the heating system to compensate for this loss and thus to maintain the given indoor temperature. The degree-days concept expresses this energy loss [6]; the energy loss per degree-day is assumed to be $\varepsilon$; this is different for each house/building and depends on the isolation of the border between indoor and outdoor with walls, windows, floor, roof, ventilation, for example. For a given time interval, the value of $tmed$ can be expressed as in equation (7). Furthermore, temperature increase energy demand ($tied$) depends on the heat energetical capacity $C$ of the house: this indicates how much energy is needed to raise the temperature by 1 degree. Therefore $tied$ is proportional to the temperature difference $\Delta T_{id}$ made and relates to the notion of capacity $C$ of the house as in the equation (8). Finally the total energy demand can be expressed as in equation (9).

$$tmed(t_1, t_2) = \int_{t_1}^{t_2} \frac{\varepsilon}{24} (T_{id} - T_{od}(t)) \, dt \hspace{1cm} (\text{assuming } T_{id} > T_{od}(t))$$  \hspace{1cm} (7)
For a small time interval with length $\Delta t$ the energy usage $eu$ is proportional to the energy demand and relates to the seasonal performance factor SPF of the heat pump as expressed in equations (10).

$$eu(t, t + \Delta t) = \frac{ed(t, t + \Delta t)}{SPF(T_{od}(t))}$$

### 13.2.2. Computational dynamic cognitive model for human action selection

Human action selection is a complex process and the exact mechanisms of it are still being explored. Nevertheless, due to the developments in brain imaging and recording techniques much knowledge about this has been developed. This knowledge provides means to design and implement models for human action selection in more accurately. Details of neurocognitive evidence behind human action formation will not be included in the current paper but can be found separately in [5,12]. Fig. 1 presents the cognitive model for action selection and Table 1 summarizes the abbreviations used. The model takes inputs from two types of world states: WS($s_k$) and WS($b_i$); here $s$ is a stimulus (that can be either external or internal to the agent) that may lead to an action execution, and $b_i$ represents the effects of the execution of an action $a_i$. The world state WS($s_k$) leads to a sensor state SS($s_k$) as input, and subsequently to a sensory representation state SR($s_k$). Moreover, the model includes both conscious and unconscious aspects. The states: SR($s_k$), SR($b_i$), PD($b_i$), PA($a_i$), F($b_i$), Pet($b_i$, $s_k$), PO($a_i$, $b_i$), and RO($a_i$, $b_i$); are considered to be unconscious and contributing to bottom-up processes, whereas the states: SD($b_i$), Att($b_i$, $s_k$), ClInt($b_i$, $s_k$), PAwr($a_i$, $b_i$, $s_k$), and RAwr($a_i$, $b_i$, $s_k$); represent more conscious influences, contributing to top-down processes. The bottom-up cognitive processes have been mapped to unconscious action formation, whereas top-down processes have been related to the conscious action formation [13–16]; it seems that the action selection process initiates from unconscious phenomena, and that later the conscious experience of this action selection is developed. The unconscious neural activations in the brain seem to be a result of habitual tasks, through the effects of prior learning, which can be automatically activated when a relevant stimulus is perceived [17]. Nevertheless, conscious awareness of action selection also plays an important role and the influence of predictive and inferential processes of action execution has been highlighted by Haggard and co-workers, providing a working mechanism for this process [13].

The unconscious bottom-up process of action selection is modelled by combining Damasio’s as if body loop [18] and James’s body loop [19]. The as if body loop presented through: PA($a_i$) $\rightarrow$ SR($b_i$) $\rightarrow$ F($b_i$) and body loop is presented
Chapter 13  337

通过 PA(\(a_i\)) → EA(\(a_i\)) → WS(b_i) → SS(b_i) → SR(b_i) → F(b_i)。根据 Damasio, 认知过程中的行为选择基于内部模拟过程，而该过程发生在行为执行之前。每个相关行为选项 PA(\(a_i\)) (刺激 \(s\) 将会有许多选项 i=1..n) 被评估 (不实际执行它们) 通过比较其与感觉相关评价值。当行为被认为是互斥的，每个准备状态 PA(\(a_i\)) 排斥所有其互补状态 PA(\(a_j\)) (当 \(j\neq i\) 时，其中 \(i\) 表示一个特定状态的一个实例)，而每个虚线紫色箭头代表直接激活到状态 \(b_i\) 从状态 \(a_i\) 而抑制所有“\(i\)th”状态在 \(b_i\) 从状态 \(a_i\)。

**Fig. 1**: Overview of the computational cognitive agent model. Here each solid black arrow represents a direct activation to state B from state A, each solid red arrow represents a direct suppression to state B from state A, each dashed red arrow represents a suppression to all the complements of ‘\(i\)th’ state on \(B_i\) from state \(A_i\) (where ‘i’ presents an instance of a particular state), and each dashed purple arrow represent a direct activation to state \(B_i\) from state \(A_i\) while suppressing all the complements of ‘\(i\)th’ state on \(B_i\) from state \(A_i\).

通过 PA(\(a_i\)) → EA(\(a_i\)) → WS(b_i) → SS(b_i) → SR(b_i) → F(b_i)。根据 Damasio, 认知过程中的行为选择基于内部模拟过程，而该过程发生在行为执行之前。每个相关行为选项 PA(\(a_i\)) (一个刺激 \(s\) 将会有许多选项 i=1..n) 被评估 (不实际执行它们) 通过比较其与感觉相关评价值。当行为被认为是互斥的，每个准备状态 PA(\(a_i\)) 为行为选项 \(a_i\) 排斥所有其互补状态 PA(\(a_j\)) (当 \(j\neq i\) 时，其中 \(i\) 表示一个特定状态的一个实例)，因此通过一种赢家通吃原则，自然地，拥有最高评价效果的感觉将执行通过身体循环。

此过程进一步增强了通过嵌入表现性欲望 \(b_i\) 和感知状态为 \(s_k\) 的 \(b_i\)。状态 PD(b_i) 促进短期利益/目标，这些目标影响行为的执行或拒绝根据其满意或不满意评价。因此，通过此状态，该代理有能力根据其欲望加强当前行为选择。
Table 1: Nomenclature for Fig. 1

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WS(W)</td>
<td>World state W (W can be either stimulus s, or effect b)</td>
</tr>
<tr>
<td>SS(W)</td>
<td>Sensor state for W</td>
</tr>
<tr>
<td>SR(W)</td>
<td>Sensory representation of W</td>
</tr>
<tr>
<td>PD(b_i)</td>
<td>Performatives desires for b_i</td>
</tr>
<tr>
<td>SD(b_i)</td>
<td>Subjective desires for b_i</td>
</tr>
<tr>
<td>PA(a_i)</td>
<td>Preparation for action a_i</td>
</tr>
<tr>
<td>Per(b_i,s_k)</td>
<td>Perception state for the effect of s_k on b_i</td>
</tr>
<tr>
<td>F(b_i)</td>
<td>Feeling for action a_i and its effects b_i</td>
</tr>
<tr>
<td>PO(a_i, b_i)</td>
<td>Prior ownership state for action a_i with b_i</td>
</tr>
<tr>
<td>Att(b_i,s_k)</td>
<td>Attention state for the effect of s_k on b_i</td>
</tr>
<tr>
<td>Clnt(b_i,s_k)</td>
<td>Conscious intention state for s_k on b_i</td>
</tr>
<tr>
<td>PAwr(a_i,b_i,s_k)</td>
<td>Prior-awareness state for action a_i with b_i and s_k</td>
</tr>
<tr>
<td>EA(a_i)</td>
<td>Execution of action a_i</td>
</tr>
<tr>
<td>RO(a_i, b_i)</td>
<td>Retrospective ownership state for action a_i with b_i</td>
</tr>
<tr>
<td>RAwr(a_i,b_i,s_k)</td>
<td>Retrospective-awareness state for action a_i with b_i and s_k</td>
</tr>
<tr>
<td>EO(a_i,b_i,s_k)</td>
<td>Communication of ownership and awareness of a_i with b_i and s_k</td>
</tr>
<tr>
<td>SR(EM_i)</td>
<td>Sensory representation of energy model</td>
</tr>
</tbody>
</table>

(through this, also some bias can be injected into the process). In parallel to the action preparation process, the perception state Per(b_i, s_k) also develops based on the salient features of the stimulus s_k; this will strengthen the bottom-up process which leads to even further strengthening of the action preparation process. Furthermore; by having a suppressive link from the Per(b_i, s_k) state to itself (see Fig. 1), the competition among perceptual entities is represented [20]. While the agent is passively (unconsciously) performing an action selection, the agent starts to activate bottom-up attention (this is represented by the link from F(b_i) to Att(b_i, s_k)). The main functionality of the bottom-up attention is to pass current information into higher order cognitive states. Due to this bottom-up attention, the agent will activate its subjective desires of b_i, which in turn leads to a conscious intention of stimulus s_k and effect b_i, and subsequently back to the attention state again. This cyclic process represents the transformation from bottom-up to top-down. Intention is considered to trigger goal directed preparation (see [12]) and therefore this model includes an effect from the conscious intention state Clnt(b_i, s_k) to the preparation state PA(a_i). This link strengthens the option of action a_i but suppresses its complementary options for all PA(a_j) with j≠i. This is also a part of the top-down process.

Once the attention (and its subjective aspects) has been developed, it injects conscious biases (through the top-down attention) into the action preparation and perception states. This is represented through the links from Att(b_i, s_k) to PA(a_i) and Per(b_i, s_k), and these links (purple dotted arrows) play a special role: while activating the matching option (i.e., the i^{th} option) they suppress all complements of the i^{th} option. The suppressive effects of this phenomenon are in line with the voluntary inhibition process (i.e., intentional suppression of an irrelevant response,
stimulus, or memory); see [21]. This emphasizes the conscious influence on action formation, and therefore attention will quickly enable the agent’s concentration, which may shorten the time required for action selection. More importantly, particular to the perceptual load, this will strengthen the current perception even further, and due to strong subjective feelings the agent may not be able to shift its attention easily. Together with these processes, the agent will develop a state of ownership, which mainly determines to what extent an agent attributes an action to himself or to another agent. Also, the agent will develop an awareness state of action $a_i$ that is related to effect $b_i$ and stimulus $s_k$. According to Haggard ([13,15]), there may be an influence from awareness states to action selection; therefore, this model includes a link from prior-awareness state $PAwr(a_i, b_i, s_k)$ to the action execution state $EA(a_i)$. The agent will execute the selected action $a_i$ and then this action will have an effect in the environment (through world state $WS(b_i)$), and be sensed again, through the body loop. Once an action is executed, this model also considers the retrospective aspects that Haggard and co-workers pointed out [13,22]. Finally, the agent has the ability to communicate the process through state $EO(a_i, b_i, s_k)$.

13.3. Cognitive driven domestic heating energy management

The main scope of this paper is to include cognitive aspects into the energy related decision making. As presented in Section 2, a mathematical model was developed earlier that can be used to calculate realistic results for energy usage of a house which uses air to water heat pump, and also a generic cognitive model was also developed earlier that mimics a detailed explanation of human action selection. In this section these two models will be combined resulting in an integrated model that will be able to simulate heating related energy management of a house together with cognitive elements. Mainly the cognitive model that was developed (see Fig. 1) is used as it is and mathematical energy model is embedded into it. Through this it is possible to keep required knowledge for the mathematical model in Section 2.1. The cognitive model accepts external information through the world state $WS(s_k)$ and this is also used as the initiation of the process. Having this external input, it triggers the agent to select a set point for house indoor temperature. The main challenge in this situation is to decide which set point to be selected and how to perform that as close as possible to human cognition involved in action selection. In a real situation, a human may utilize his or her desires on comfort level, perception on energy related savings, and intentions to save money or energy. An agent can prepare for many possible options through multiple action preparation states $PA(a_i)$. For example, this can be interpreted as selecting set point value to $18^\circ C$ ($PA(a_1)$), $16^\circ C$ ($PA(a_2)$), or $14^\circ C$ ($PA(a_3)$). Before actually executing any option, the agent internally simulates the effects of these prepared options individually through the as
if body loop. By comparing the feeling-related valuations associated to their individual effects that contribute to select the winning option: the option that has the strongest positive feeling will be the winning option to execute. To facilitate this it is required to have different weight values for the prediction connection from PA(a_i) to SR(b_i). Having a strong link contributes to a strong feeling whereas having a weak link contributes to a poor feeling. In this model there are two types of connection links available: activation and suppressive. Activation links hold positive weight values (max value is +1 and the minimum is 0) and a suppressive link holds a negative weight value (max value is 0 and the minimum is -1). The complete information of the chosen weight values in Fig. 1 is presented in Table 2; the challenge in here is to select appropriate realistic values for them.

Through the as if body loop, by having different values for the links from PA(a_i) to SR(b_i) the agent develops different feelings about the options. Having significant differences between these values naturally leads to selection of the candidate with the strongest weight value. This can be considered as a habitual situation, in which user has executed the same decision many times and that has been hard coded by a very strong strength for that link (a form of neural plasticity). To eliminate such a bias effect it is made sure that the differences between weight values for the three

<table>
<thead>
<tr>
<th>from states</th>
<th>to state</th>
<th>weights</th>
<th>L.P</th>
</tr>
</thead>
<tbody>
<tr>
<td>WS(s_k)</td>
<td>SS(s_k)</td>
<td>ω1</td>
<td>LP1</td>
</tr>
<tr>
<td>SS(s_k)</td>
<td>SR(s_k)</td>
<td>ω2</td>
<td>LP2</td>
</tr>
<tr>
<td>SR(s_k), Att'(b_i, s_k)</td>
<td>PD(b_i)</td>
<td>ω3, ω4</td>
<td>LP3</td>
</tr>
<tr>
<td>SR(s_k), PD(b_i), F(b_i), Per(b_i, s_k), Clnt(b_i, s_k), Clnt'(b_i, s_k), Att(b_i, s_k), Att'(b_i, s_k), PA(a_j) (j≠i)</td>
<td>Per(b_i, s_k)</td>
<td>ω14, ω15, ω16, ω17, ω18, ω19, ω20, ω21</td>
<td>LP5</td>
</tr>
<tr>
<td>SR(s_k), PD(b_i), SD(b_i), Att(b_i, s_k), Att'(b_i, s_k), PAwr(a_j, b_j, s_k), Per'(b_i, s_k), SR(EM_i)</td>
<td>SD(b_i)</td>
<td>ω22, ω23</td>
<td>LP6</td>
</tr>
<tr>
<td>PA(a_i), SS(b_i), PO(a_j, b_j)</td>
<td>F(b_i)</td>
<td>ω24, ω25</td>
<td>LP7</td>
</tr>
<tr>
<td>SR(b_i), PD(b_i)</td>
<td>Clnt(b_i, s_k), PAwr(a_j, b_j, s_k)</td>
<td>Att(b_i, s_k)</td>
<td>ω26, ω27, ω28</td>
</tr>
<tr>
<td>SR(s_k), Att(b_i, s_k), Clnt'(b_i, s_k), SR(EM_i)</td>
<td>Clnt(b_i, s_k)</td>
<td>ω29, ω30, ω31, ω32, ω33, ω34</td>
<td>LP9</td>
</tr>
<tr>
<td>Per(b_i, s_k), SD(b_i), SR(EM_i)</td>
<td>PO(a_j, b_j)</td>
<td>ω35, ω36, ω37, ω38, ω39, ω40, ω41</td>
<td>LP10</td>
</tr>
<tr>
<td>F(b_i), PA(a_i), RO(a_j, b_j)</td>
<td>PO(a_j, b_j)</td>
<td>ω34, ω35, ω36</td>
<td>LP11</td>
</tr>
<tr>
<td>PO(a_j, b_j), F(b_i), Att(b_i, s_k), Clnt(b_i, s_k), RAwr(a_j, b_j, s_k)</td>
<td>PAwr(a_j, b_j, s_k)</td>
<td>ω42, ω43, ω44</td>
<td>LP12</td>
</tr>
<tr>
<td>PA(a_i), PO(a_j, b_j), PAwr(a_j, b_j, s_k)</td>
<td>EA(a_i)</td>
<td>ω42, ω43, ω44</td>
<td>LP13</td>
</tr>
<tr>
<td>EA(a_i)</td>
<td>WS(b_i)</td>
<td>ω45</td>
<td>LP14</td>
</tr>
<tr>
<td>WS(b_i)</td>
<td>SS(b_i)</td>
<td>ω46</td>
<td>LP15</td>
</tr>
<tr>
<td>EA(a_i), F(b_i), PO(a_j, b_j)</td>
<td>RO(a_j, b_j)</td>
<td>ω47, ω48, ω49</td>
<td>LP16</td>
</tr>
<tr>
<td>RO(a_j, b_j), F(b_i), PAwr(a_j, b_j, s_k)</td>
<td>RAwr(a_j, b_j, s_k)</td>
<td>ω50, ω51, ω52</td>
<td>LP17</td>
</tr>
<tr>
<td>RO(a_i, b_j), Clnt(b_i, s_k), RAwr(a_i, b_j, s_k)</td>
<td>EO(a_i, b_j, s_k)</td>
<td>ω53, ω54, ω55</td>
<td>LP18</td>
</tr>
</tbody>
</table>
options in the links from $PA(a_i)$ to $SR(b_j)$ are not significant (the selected difference is 0.1). Having such a situation now it is difficult for the agent to select its option and further cognitive aspects needs to be considered. For this purpose performative desires states $PD(b_j)$ are used. These states facilitate short-term interests/goals that influence either selecting or rejecting an action due to its satisfactory or less satisfactory valuation (for more information see [23]). This can be used to express a human’s expectation/desires on comfort level. If someone is more serious about its comfort level, it is natural to bias set points towards high temperature (considering a heating situation). Performative desire states $PD(b_j)$ directly affect feeling and action preparation states. Therefore, having a strong desire on comfort level will inject more bias to higher temperature set point values. This contributes to select an option easily through a strong predictive feeling and as the human is mainly (or partly) in unconscious mode or in a habitual mode this should quickly lead to action selection.

In parallel to the above process a perception state $Per(b_j, s_k)$ also plays a key role. Having a strong perception leads to select an action quickly and contributes to strong emotions (for more details about the effects of perception on cognitive process see [5,24]). It is possible to easily develop a wrong or bias perception immediately from the input stimuli. Having a strong bias perception, may quickly lead to selection of a wrong option and the agent may be unable to correct this. This is considered to be the most frequent problem for poor situation awareness [25,26]. This model also explains the effects of bias perception, and specifically it has effects on the action preparation states $PA(a_i)$ and conscious intention states $CInt(b_j, s_k)$. Therefore, having a strong perception injects strong influence on action preparation and leads to strengthening of the feeling through the as if body loop. The agent develops perceptions based on the results of the energy model. It is a well known fact that the energy cost that we save for heating by reducing the set point for one or a few degrees is not huge. If the agent has a biased perception, he or she will calculate the energy that will be saved by setting those set points and if the savings are not attractive, a strong perception on the highest set point value will be developed.

The effects and processes discussed above are mainly contributing to the unconscious or habitual form of action selection. Nevertheless, conscious aspects are also very important and especially if the human is expecting something contradicting with his/her comfort, this is a vital part. As highlighted in Section 2.2, this model includes both unconscious and conscious processes. The initiation of the process is always as explained in above paragraphs. Once the agent developed predictive feeling about each option, that will lead to activation of the agent’s attention (for detail cognitive basis of this see [5,12]). This is referred to as bottom-up attention and through this the agent starts to get more attention on what is
happening unconsciously so far. Having developed some attention, this will further affect subjective desires for \( b_i \). This is the state which represents long term goals of the user. The activation level of the subjective desire states depends on the results provided by the mathematical model. Due to triggering the mathematical model in here, now it is able to predict the energy required for individual situations/options. By having different energy needs for different set points, now the agent can decide which is the most economical (assuming the subjective desires to save money and energy). Therefore, the set point which is having the lowest energy demand should be highly motivated in this regard. As a result of that, the state SD\( (b_i) \) (for the 14°C option) strongly affects the conscious intention state for \( s_k \) on \( b_i \). Therefore, for the lowest energy demand CInt\( (b_i, s_k) \) should be high and for the highest energy demand this should be relatively low. Having this feature now the state of CInt\( (b_i, s_k) \) becomes strong only if the energy used is minimum. Therefore, the agent naturally develops a strong intention to select that option while other options are having a weak intention. There is a cyclic process among Att\( (b_i, s_k) \) to SD\( (b_i) \) to CInt\( (b_i, s_k) \). Through this process it is natural that over time the agent experiences strong attention and intention on the option that minimizes the energy usage (together with the subjective desire of saving energy and money). Having this strong attention, now a new effect called top-down attention activates. Top-down attention involves a link from Att\( (b_i, s_k) \) to PA\( (a_i) \). This is a very strong positive link and injects strong bias on action preparation consciously. Furthermore while attention for a particular option is strengthening its action preparation suppresses the preparations PA\( (a_j) \) for all the complements of that option (those with \( j \neq i \)). Due to this, when a strong attention developed for an option, it will dominate the action selection process and other options will be diluted over time. This can be used to demonstrate the interplay between conscious and unconscious processes. Furthermore, in line with these processes, the agent develops a prior awareness state. This state also affects attention and also through this attention gets stronger. Finally, the state PAwr\( (a_i, b_i, s_k) \) affects action execution also, and this contributes to reduce the execution time under the conscious mode (conscious mode is relatively slow when compared with pure unconscious mode).

These processes will lead the agent to select an option and execute it through the state EA\( (a_j) \). Once the action is executed the feelings of the actual action will be combined with the predicted feeling. Nevertheless, these retrospective effects are out of the scope of the current work; more detailed information on these processes can be found on [27]. These retrospective effects are very useful for learning and corrections if there is a mismatch between what is predicted and what is observed once the action is executed (see [5]).
13.3.1. Model Compilation as a Dynamic System

The above mentioned approach has been compiled into an executable model. For this a dynamical systems perspective is used, as explained in [28]. Each connection between states in Fig. 1 (the ones specified in Table 2) has been given a weight value (where $\omega_{ji}$ represents the weight of the connection from state j to i) that varies between 1 and -1 as indicated in Table 2. Weight values are non-negative in general, except if they represent a suppressive (or inhibiting) link (see caption of Fig. 1). In addition to the link weights, each state includes an additional parameter called speed factor $\gamma_i$, indicating the speed by which the activation level of the state ‘i’ is updated upon receiving input from other states. Two different speed factor values are used, namely fast and slow values: fast values are used for internal states and slow values for external states (i.e., for WS(W), SS(W), EA($a_i$), and EO($a_i, b_i, s_j$)). The level of activation of a state depends on multiple other states that are directly attached toward it. Therefore, incoming activation levels from other states are combined to some aggregated input and affect the current activation level according to differential equation (11). As the combination function for each state, a continuous logistic threshold function is used: see equation (12), where $\sigma$ is the steepness, and $\tau$ the threshold value. When the aggregated input is negative, equation (13) is used. To achieve the desired temporal behaviour of each state as a dynamical system, the difference equation represented by equation (14) is used (where $\Delta t$ is the time step size).

\[
\frac{dy_i}{dt} = \gamma_i \left[ f \left( \sigma, \tau, \sum_j \omega_{ji} y_j \right) - y_i \right]
\]

\[
f(\sigma, \tau, X) = \left( \frac{1}{1 + e^{-\sigma(X-\tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{-\sigma \tau}) \quad \text{when } X > 0
\]

\[
f(\sigma, \tau, X) = 0; \quad \text{when } X \leq 0
\]

\[
y_i(t + \Delta t) = y_i(t) + \gamma_i \left[ f \left( \sigma, \tau, \sum_{j \in s(i)} \omega_{ji} y_j \right) - y_i(t) \right] \Delta t
\]

In addition to this setup the mathematical energy model is integrated within the process and it triggers automatically and is represented as an internal state. Outputs of the mathematical model come from equation (10) but it is coupled with equations (1 to 9). Therefore, internally all these equations execute to represent the situation in the environment (both external and internal) to retrieve energy usage values for prepared set points. The generated energy usages values are represented in the cognitive model as sensory representation states (see Fig. 1, state SR(EMi)). This cognitive model assumes that the maximum strength of a state is +1 and the
minimum is 0. Therefore, it is required to have a transformation to convert actual energy usage values into the scale of 0 to +1. Equation (15) is used for this requirement, where it is assumed to be that EM\text{min}, and EM\text{max} are minimum and maximum energy usage values provided by the mathematical model under a given situation.

\[
y_{SR,EM,i,T^e}(t) = 1 - \left( \frac{x_i - EM_{min}}{EM_{max} - EM_{min}} \right) \text{ where, max} \\
= \left( EM_{max} + \frac{EM_{max} - EM_{min}}{2} \right) (15)
\]

At each step (with step size \(\Delta t\)) all the other states are updated (through equation 14) and the results of the mathematical model are used for the states Per\((b, s)\), SD\((b, s)\), and Clnt\((b, s)\). The effects of the mathematical model influence the behaviour of these three states is explained in equations (16, 17, and 18). Let’s assume that SR\((EM_i)\) values provided by simulation of the mathematical model for night goal temperatures: 18°C, 16°C, and 14°C are \(\alpha, \beta, \) and \(\delta\) respectively.

\[
y_{SD,T^e}(t + \Delta t) = y_{SD,T^e}(t) \\
+ y_{SD} \left[ f \left( \sigma, \tau, \omega_{29,T^e} y_{SR,T^e}(t) + \omega_{30,T^e} y_{Att,T^e}(t) \\
+ \omega_{31,T^e} y_{Clnt,T^e}(t) + \omega_{em2,T^e} y_{SR,EM_i,T^e}(t) \right) \\
- y_{SD,T^e}(t) \right] \Delta t \text{ here if } \alpha > \beta > \delta \text{ then } \omega_{em2,18^c} < \omega_{em2,16^c} < \omega_{em2,14^c} \tag{16}
\]

\[
y_{Clnt,T^e}(t + \Delta t) = y_{Clnt,T^e}(t) \\
+ y_{Clnt} \left[ f \left( \sigma, \tau, \omega_{32,T^e} y_{Per,T^e}(t) + \omega_{33,T^e} y_{SD,T^e}(t) \\
+ \omega_{em3,T^e} y_{SR,EM_i,T^e}(t) - y_{Clnt,T^e}(t) \right) \Delta t \text{ here if } \alpha > \beta > \delta \text{ then } \omega_{em3,18^c} < \omega_{em3,16^c} < \omega_{em3,14^c} \tag{17}
\]

\[
y_{Per,T^e}(t + \Delta t) = y_{Per,T^e}(t) \\
+ y_{Per} \left[ f \left( \sigma, \tau, \omega_{14,T^e} y_{SR,T^e}(t) + \omega_{15,T^e} y_{PD,T^e}(t) \\
+ \omega_{16,T^e} y_{SD,T^e}(t) + \omega_{17,T^e} y_{Att,T^e}(t) + \omega_{18,T^e} y_{Att',T^e}(t) \\
+ \omega_{19,T^e} y_{PAWR,T^e}(t) + \omega_{20,T^e} y_{Per',T^e}(t) + \omega_{em3,T^e} y_{SR,EM_i,T^e}(t) \right) \\
- y_{Per,T^e}(t) \right] \Delta t \text{ here, if } (\alpha - \beta < \delta) < \theta \text{ for example } \theta = 1.0 \text{ then } (\omega_{14,18^c}, \omega_{15,18^c}, \omega_{16,18^c}, \omega_{17,18^c}) \\
> (\omega_{14,16^c}, \omega_{15,16^c}, \omega_{16,16^c}, \omega_{17,16^c}) \\
> (\omega_{14,14^c}, \omega_{15,14^c}, \omega_{16,14^c}, \omega_{17,14^c}) \tag{18}
\]
In addition, parameter estimation is the difficult task in cognitive modelling; the parameter values used in [5] are used with very few changes (see next Section for information).

13.4. Integrated model validation through simulations

Having identified an integrated model of combined cognitive and energy related processes it is necessary to explore its behaviour for practical situations. Four scenarios are selected for simulations that cover the main aspects of the integrated model. For all the scenarios a generic setup is selected. The behaviour was analysed specifically for data available from indoor temperature 20°C at time 21:00hrs February 1, 2012 to 06:00hrs February 2, 2012 (the same data set used in [3]). It is assumed that the heating program is not using energy until the temperature reaches the night goal temperature $T_{ng}$ (until then autonomous cooling down takes place). Once the indoor temperature becomes $T_{ng}$ that temperature is maintained until 06:00hrs 2nd February 2012, and the indoor temperature is increased from $T_{ng}$ to 20°C at 06:00hrs. Throughout this time interval, the outdoor temperature is assumed to be behaving as in equation (2) and SPF is calculated as in the equation (6). According to the collected data in [29] for this period of time, the minimum temperature $T_{min}=-8.8^\circ C$, the temperature at sunset $T_{sunset}=-2.52^\circ C$, the time of sunset $t_{sunset}=17:00$, and the outdoor temperature at 21:00hrs $T_{od}=-6.6^\circ C$. Furthermore, for the remaining parameters the values were: $C=4.6$, $\varepsilon=4$, $\alpha=0.25$, $d=0.1$, $T_w=50^\circ C$, and time step size $\Delta t = 6$min was for the simulation. Furthermore only 3 night goal temperatures ($T_{ng}$) were selected 18°C, 16°C, and 14°C.

The first scenario explains a situation where the agent is having a strong performative desire for the higher comfort level and a less strong subjective desire to save money and energy. Therefore, this higher night goal temperature is expected to be selected. The second scenario explains a situation where the agent has a strong subjective desire to save money and energy and a less strong performative desire for comfort. Therefore the night goal temperature with the lowest value is expected to be selected. The third scenario illustrates a situation where the agent has a wrong or biased perception. Therefore, the agent has the perception that if the energy usage is not significant for different night goal temperatures then there is no reason to select a low night goal temperature. The fourth scenario explains a situation where the agent has both a relatively high performative desire and subjective desire. This should demonstrate the compromising effect of these desires and the agent may be expected to select an average value for night goal temperature. This system is implemented as specified in Section 3.1 and implemented by using Java and
provided all parameter values in XML form. The detail information about parameter values of each scenario can be found separately\(^2\).

### 13.4.1. Scenario 1: Agent with a strong performative desire for higher comfort level

In this situation it is assumed to be that agent has the impression a good comfort level throughout the night period is required and it is believed that a higher night goal temperature will fulfil this. As explained in Section 3 the agent initiates the process with external stimuli and prepares for three options: 18°C, 16°C, and 14°C. For the weights of the as if body loop from PA\((a_i)\) to SR\((b_i)\) link only small differences are given for the different options (0.9, 0.8, and 0.7 respectively, and this is same for all the scenarios). In addition to this difference all the other parameter values of the cognitive model are identical except for the state PD\((b_i)\). To facilitate a strong performative desire for a higher comfort level, different weight values were used for the connection from SR\((s_i)\) to PD\((b_i)\), such that PD\((b_i)\) becomes activated stronger for higher goal temperature values. Therefore for the SR\((s_i)\) to PD\((b_i)\) link values 0.9, 0.6, and 0.3 are used, respectively. Fig. 2 shows the simulation results.

It is very clear that having identical parameter values for all the options except for the two links, the agent has cognitively selected the highest night goal temperature. Furthermore, the results of the mathematical energy model do not directly affect performative desires states and weak values were used for \(\omega_{em1}\), \(\omega_{em2}\), and \(\omega_{em3}\). For each scenario three simulation graphs were generated and only one option is executed and only that is presented in this paper. Complete simulations

---

\(^2\) http://www.few.vu.nl/~dte220/IEREK15ParameterData.zip
results for 3 options are separately included in an external appendix\(^3\) (for all the scenarios). Therefore it is clear that the agent has developed strong activations only for the highest set point value (18°C); for the other two options no sufficient strength has developed to execute the action. Furthermore, the emerging orders of activation of the states are also very important to validate the model behaviour and results are fully complied with the expected trace order in [5]. At the beginning PD\((b_i)\) for 18°C option has shown a good strength while the same for 16°C is slightly weak and 14°C is very weak. Through this other processes have affected (specially the as if body loop) and naturally the predictive feeling of the 18°C option becomes strong and the agent executes it.

13.4.2. Scenario 2: Agent with a strong subjective desire to save money and energy

This scenario presents effects of having a strong subjective desire to save money and energy and a less strong desire for comfort. In this situation the agent has a moderate value for performative desires and it is not too strong for 18°C option. For the connections from SR\((s_i)\) to PD\((b_i)\) the same value (0.6) is given for each option. Therefore, no bias is created through performative desires states. The agent uses the mathematical energy model to generate the activation level for the state SD\((b_i)\). From the predicted data of the mathematical energy model they may require 23.6 kWh, 22.71 kWh, 22.12 kWh for the 18°C, 16°C, and 14°C option, respectively. Therefore, this is used to decide how strong the weight value of \(\omega_{em2}\): for the highest energy demand value 0.3, for the average energy demand value 0.5, and for the lowest 0.9 is assigned for options 18°C, 16°C, and 14°C respectively. Therefore, the agent develops a strong subjective desire for the option 14°C.

In addition to this, relative to the energy demand of each option agent set the strength of the weights \(\omega_{em3}\) (for each option). Therefore, for the \(\omega_{em2}\) on 18°C get the value 0.3, 16°C get the value 0.5, and 14°C get the value 0.9. Having this configuration with identical values for other weights (for each option separately) agent has able to successfully select the set point value of 14°C. This is what is expected and Fig. 3 provides the details particular to this option (simulation results of other two options can be found in the external appendix). Therefore, this model has successfully integrated the predictions of energy model into the cognitive model and provided realistic results. In simulation results it is clear that at the beginning the same level of PD\((b_i)\) state has activated for 3 options and agent has developed almost the same predictive feeling value for those options. Therefore, bottom-up process passes this information to the higher cognitive levels and agent experiences attention on these three options (see external appendix). Nevertheless with the

\(^3\) http://www.few.vu.nl/~dte220/IEREK15ExternalAppendix.docx
Fig. 3. Agent with a strong subjective desire to save money and energy.

effects of SD\( (b_i) \) state only the CInt\( (b_i, s_k) \) for option 18°C has strongly developed and this has reflected through the top-down attention. This top-down attention effects on PA\( (a_i) \) on option 18°C and therefore, predictive feeling of this has increased significantly (while the top-attention of option 18°C is suppressing the other options). Therefore, agent leads to selection 18°C option and it has successfully executed (see Fig. 3).

13.4.3. Scenario 3: Agent with a wrong or bias perception

This is a very important situation for many domains. According to Endsley’s model for situation awareness [25] bias or wrong perception is considered to be as the main problem for Level 1 situation awareness and it is the cause for 76% errors in the aviation domain [26]. Here, agent is with the impression that if the energy demand for each option is not significant then there is no reason to compromise comfort level selecting a low night goal temperature. Agent prepares the process with an external stimulus and prepare for three options. In this configuration all the parameters are the same for 3 options except for predictive link in the as if body loop and for the state perception. Therefore, initially agent has developed the same strength for PD\( (b_i) \) but with the activation of Per\( (b_i, s_k) \) things have changed. Simulation results of this are presented in Fig. 4 (for all the options see the external appendix).

The integrated model decides the strength of each perception state (for each option) related to the energy usage values provided by the mathematical model. The mathematical energy model calculates the energy demand for 3 options separately and that holds the values: 23.6 kWh, 22.71 kWh, 22.12 kWh for 18°C, 16°C, and 14°C. It is very clear that this difference is not very significant for each option and there is no significant energy saving by selecting a low night goal set point value.
Therefore the agent develops a strong perception only for the 18°C option. This achieved by providing different weight values for $\omega_{em1}$ for each option. The values 1.0, 0.3, and 0.2 are used for the 18°C, 16°C, and 14°C options, respectively. Due to this biased or wrong perception it positively strongly effect on PA($a_i$) state on option 18°C. Having a self-suppressive link on state PA($a_i$) this will suppresses all its complements. Therefore As PA($a_i$) for option 18°C is the strongest it suppresses all its complements and naturally strengthen the feeling of predictive effect while for the other two options this is get diluted. Finally agent leads to execute to select option 18°C.

13.4.4. Scenario 4: Agent with a strong performative desire and subjective desire

This is a special situation where agent expects both performative and subjective desires. This may seems like a conflict but in the reality this is a common situation. In such a situation it may not possible to only consider the effects of performative desire or the effects of subjective desire. Therefore combination of these two should reflect in the results and it will be a compromise of these two. According to the available three options it is expected to select 16°C options as it has sufficient comfort level and energy saving. Agent initiates the model with an external stimulus. In this situation for the predictive link in the as if body loop (i.e., from PA($a_i$) to SR($b_i$)) the same values are used as in the scenario 1. In addition to that for the link from SR($s_i$) to PD($b_i$) different values are used: 0.9, 0.9, and 0.6 for options 18°C, 16°C, and 14°C respectively. Therefore it is clear that for both 18°C, 16°C the same performative desire strength can be expected. Therefore, as explained for the scenario 1 now 18°C option will not dominate the process. Similar to the other 3 simulations, for this case also the model uses the mathematical model.
to incorporate the expected energy usage. For the 14°C option it is smallest and therefore for the weight $\omega_{em2}$ (for option 14°C) the value 0.8 is provided, for the 16°C option a slightly higher value (0.7) is assigned for the highest energy usage option which is 18°C a smaller value (0.3) is assigned. This also confirms that there is no strong bias through 18°C, 16°C options. In this situation by considering the energy demand values agent assigns values for the weights $\omega_{em3}$ as 0.6, 0.9, 0.6 for three options respectively. As the state $SD(b_i)$ is affected from $Att(b_i, s_k)$ it has information about unconscious processes and therefore by considering the required energy demands a slight high value is given for the option 16°C while for the other two the same value is assigned. Having this configuration agent has developed strong attention on 16°C option due to the cyclic loop $SD(b_i) \rightarrow CInt(b_i, s_k) \rightarrow Att(b_i, s_k)$. Top-down attention effect enables and it has positively effected on action preparation of 16°C option (while it is suppressing the other two options’ action preparations). Therefore, agent is experiencing a good predictive effect on this option and finally leads to execute it. Fig. 5 shows the simulation results for the option 16°C.

13.5. Discussion

Human behaviour is not always easy to predict and may be complex. This is even more so if the environment in which the human functions is complex and dynamic. One example of such a complex and dynamic environment is domestic heating with the dynamics associated to indoor and outdoor air temperatures, required comfortable temperature set points over time, parameters of the heating source and system, and energy loss rate and capacity of a house. It is difficult to conduct real world experiments to analyse the dynamics and optimal efficiency of a heating system in actual daily use under all circumstances and constraints. As an
alternative a simulation-based analysis may be an alternative. This paper has presented a simulation-based analysis of a domestic heating energy management system using an air to water heat pump. It has integrated an earlier developed mathematical model for this heating system’s performance, and a computational dynamic cognitive model for the human’s behaviour which was developed based on evidence from Cognitive Neuroscience. The model covers performative and subjective desires, perception, emotion, feeling, ownership, attention, intention, and awareness. This model has an adequate level of detail to be used as a coherent basis for various experimental needs to analyse behaviours. The presented approach provides an experimental workbench in which complex scenarios can be explored by simulation experiments.

It is important to be able to examine the choices of human behaviour on energy usage with a realistic setup as presented in this paper. Through this it gives an overall idea on possible impacts and allows additional information on how to motivate persons or correct their biased perceptions in a methodological approach to support lifestyle and lifestyle change in relation to energy management. This paper provides a contribution to this line of research, and may be useful as a tool for learning purposes in order to improve the awareness of humans on (heating) energy management.

References

Part V:

Discussion

This thesis is composed of a collection of published research papers (except one) that collectively contribute to a research question how dynamic computational cognitive models for human action selection can be designed, developed, simulated and applied, and what is the role of awareness and cognitive control in such models. Each previous chapter provides detailed information on specific (sub) research questions as stated in the introduction chapter. This part summarises the research presented and highlights implications of this work. Furthermore, an outline of some future work is included providing links to new interesting challenges to explore.
Chapter 14: Discussion

This thesis explores processes involved in human awareness and related control of behaviour which have received much interest within the research community. The fundamental research question of this thesis is how we can understand and explain human awareness and control in behavioural choices through dynamic computational models. Various approaches have been proposed to understand and explain the working mechanisms of human cognition underlying such choices. For this purpose relatively new knowledge from the neuro-cognitive research area (strongly fuelled by neuro-imaging and recording techniques) and a dynamic system modelling approach have been used. Furthermore, the explored dynamic cognitive models have been validated for socio-technical systems in aviation and energy domains.

The main research question was scrutinised through four main research questions. Furthermore, those four research questions were also refined into a few more atomic research questions, each of which is more manageable and sufficiently specific to explore. Following the breakdown of research questions of this thesis, once the obtained answers are integrated, this leads to answers to the main question ‘how can dynamic computational cognitive models for human action selection be designed, developed, simulated and applied and what is the role of awareness and cognitive control in such models’?

1) What is the role of awareness in human action selection?
   a) What is the role of unconscious and conscious processes and their interaction in human cognition for action selection?
   b) How does the internal prediction process shape or contribute to the (prior) awareness of the action?
   c) How does the inferential sense making shape or contribute to the (retrospective) awareness of the action execution?
   d) How does the awareness contribute to action execution?

2) What are the roles of emotions, ownership and cognitive control in human action selection?
   a) What are the roles of action ownership and action awareness states in action selection?
   b) What can be learnt from cognitive/affective/behavioural sciences on how perception, attention, intention, emotion and awareness contribute to action formation and cognitive control?
   c) What can be learnt from cognitive/affective/behavioural sciences on how an emotion generation process interacts with action selection?
d) How to model the interplay between conscious vs. unconscious processes in human action formation?

3) How to design cognitive models for joint decision making on action selection and what is the role of cognitive metaphors in such models?
   a) What can be learnt from Buddhist explanations for analogy making?
   b) What can be learnt from cognitive metaphors for joint decision making on action selection?

4) How can the computational models be applied in complex real-world domains?
   a) How to estimate parameters in dynamic cognitive models, especially given incomplete empirical features or patterns over time?
   b) How to embed nature-inspired human cognitive processes for situation awareness in critical domains?
   c) How to improve the analysis of domestic energy management for heating through dynamic computational models?
   d) How to utilize cognitive models in intelligent energy management through simulations to uplift the state of art in current system automation?

14.1. Summary of research questions

The main focus of the thesis is on computational cognitive modelling for human action selection and what is the role of awareness and cognitive control in such models. Cognitive, behavioural and affective science related research was explored in the direction of the above theme. The human brain is a complex, intricate, adaptive, dynamical system; it is difficult to unravel it and comprehend its mechanisms (Bressler, 1995; Fries, 2005; Friston, 2002; McIntosh, 2000; Mesulam, 1998; Swanson, 2012). The rapid developments in brain imaging and recording techniques (especially in the last three decades) allows the research community to uplift the understanding of brain processes and mechanisms while forming new branches of research to scrutinize these functionalities from different points of view (Posner & Raichle, 1999; Raichle, 2003; Kwong et al., 1992). This literature provides much detail on very specific micro-level information about certain processes. One challenge in the work described here was to integrate such interesting micro-level information of cognitive processes into one compound model. The processes under investigation include various abstracted concepts and relations between them. As part of the abstraction process, key important concepts (e.g., awareness, emotion, ownership, perception, attention, intention, desire, and metaphors) were explored through the literature on how these concepts are getting
affected and how they in turn affect other concepts. These concepts have been represented as states in the designed cognitive models and the links between these states as arcs between states in the model. Having a cognitive model doesn’t represent any value, unless it is validated in one way or the other. Cognitive model validation is another challenge, mainly due to the lack of information on complete empirical sequences over time, the many parameters representing individual characteristics, and difficulties with parameter estimation approaches for such large numbers of parameters. This specific challenge was addressed through integrating different expected features or patterns over time, as known from literature. Having information on such expected patterns, two parameter estimation approaches were used to fine-tune cognitive models to their expected behaviour. In this way, by identifying many interesting scenarios, the designed and developed models were validated. Furthermore, cognitive models have been validated for their application on socio-technical systems in aviation and energy domains. A summary of how each (sub) research question is reflected in the chapters of the thesis is as follows.

<table>
<thead>
<tr>
<th>Part</th>
<th>Chapter</th>
<th>Research Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>2</td>
<td>1.a, 1.b, 1.c, 1.d, 2.a</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.a, 1.b, 1.d</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.d, 2.a, 2.c, 2.d</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.a, 2.b, 2.d</td>
</tr>
<tr>
<td>III</td>
<td>6</td>
<td>3.a</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>3.b</td>
</tr>
<tr>
<td>IV</td>
<td>8</td>
<td>4.a</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.c, 1.d, 2.a, 2.b, 4.b</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>4.c</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>4.c</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>4.c</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>4.d</td>
</tr>
</tbody>
</table>

14.1.1. What is the role of awareness in human action selection?

Among the many processes in the human brain, human awareness has received much attention from many researchers; e.g., (Baumeister, Masicampo, & Vohs, 2011; Desantis, Hughes, & Waszak, 2012; D’Ostilio & Garraux, 2012; Haggard, Clark, & Kalogeras, 2002; Haynes, 2011; Libet, Gleason, Wright, & Pearl, 1983; J. W. Moore, Lagnado, Deal, & Haggard, 2009; Moore & Obhi, 2012; Walsh & Haggard, 2013; Wegner, 2002). Among the many relations between human awareness and cognitive functionalities, action selection has obtained strong recognition within the research community (Baumeister, Masicampo, & Vohs, 2011; Haggard, Clark, & Kalogeras, 2002; Libet, Gleason, Wright, & Pearl, 1983; Walsh & Haggard, 2013). It is a common belief among humans that we are aware of what we are doing and therefore we are responsible for our actions. However, it is
still a question how we actually select our actions and why we select that action in a
given situation instead of something else (especially in complex situations). Is that a
purely random selection? If not, how do we make a rational choice for an action and
how does awareness contribute to it, if it does?

When studying the role of awareness in human action selection unconscious and
conscious processes both can play an important role. What are these roles of
unconscious and conscious processes and their interaction in human cognition for
action selection, is explored in this thesis. Chapters 2, 3, and 5 provides detailed
information in particular on this question. The research covered addresses and
integrates two hypotheses:

(1) Awareness of action selection is not directly causing action execution (or
behaviour) but comes afterwards, as an effect of unconscious processes of
action preparation (Baumeister, Masicampo, & Vohs, 2011; D'Ostilio &
Garraux, 2012; Haynes, 2011; Libet, Gleason, Wright, & Pearl, 1983;
Wegner, 2002)

(2) Both predictive and inferential processes related to the action preparation
and execution may contribute to the conscious awareness of the action,
making awareness of an action a dynamic combination of both prior
awareness (through predictive motor control processes) and retrospective
awareness (through inferential sense-making processes) relative to the
action execution (Desantis, Hughes, & Waszak, 2012; Haggard, Clark, &
Kalogeras, 2002; Moore, Lagnado, Deal, & Haggard, 2009; Moore &
Obhi, 2012; Walsh & Haggard, 2013).

This combination leads to integration of both conscious and unconscious
explanations for action awareness and ownership, and facilitates a working process
for action awareness through the interplay between conscious and unconscious
processes.

The cognitive aspects before and after action execution have been addressed in
literature such as (Haggard, Clark, & Kalogeras, 2002; Moore, Lagnado, Deal, &
Haggard, 2009; Moore & Obhi, 2012; Walsh & Haggard, 2013). In this research, a
distinction is made between two (sub) research questions: how does the internal
prediction process shape or contribute to the (prior) awareness of the action and
how does the inferential sense making process shape or contribute to the
(retrospective) awareness of the action execution? Chapter 2 and 3 provide details
for the first question and Chapters 2 and 9 contribute to the second. Predictive
processes related to action preparation and execution may contribute to the
conscious awareness of the action. This form of awareness is referred to as prior
awareness. In this regard internal prediction processes (internal simulation of the
action) play a key role in shaping or at least contributing to prior awareness.
Research by Haggard and co-workers provides much evidence for this issue: (Brass
One specific occurrence of this prediction process is based on Damasio’s ‘as if’ body loops (Damasio, 1999, 2005, 2012), which predict particular body states. Inferential sense-making processes play an important role in relation to the effects after action execution; this leads to the development of retrospective awareness (Desantis, Hughes, & Waszak, 2012; Haggard, Clark, & Kalogeras, 2002; Moore, Lagnado, Deal, & Haggard, 2009; Moore & Obhi, 2012; Walsh & Haggard, 2013). Retrospective awareness answers the question: ‘What have I done?’. Such a retrospective awareness state often relates to acknowledging others and taking responsibility for having performed the action. It may also play an important role in learning based on experience: evaluating the obtained effect in a conscious manner leads to an improvement of the performance of the action selection in the future. By combining this internal prediction process and inferential sense making process through awareness states, a person can examine the performance of an executed action, which is very important, in particular to learning and to functioning as a social being.

One more important sub-research question is how the awareness contributes to the actual execution of an action? When a prior awareness state occurs, a person may have become aware of going to perform the action. Having such a prior awareness state may leave open the question whether the agent is able to consciously decide to perform or not to perform the action (Baumeister, Masicampo, & Vohs, 2011; Custers & Aarts, 2010; D’Ostilio & Garraux, 2012; Haynes, 2011; Libet, Gleason, Wright, & Pearl, 1983; Wegner, 2002). Through bottom-up activations, the agent develops prior awareness and through this the agent can inject some bias to the current unconscious processes. This may strengthen a weaker action option and improve the predictive feeling of that option, which may lead to getting it executed. Furthermore, the prior awareness can also directly strengthen the action execution state too. Another aspect to look at in the context of this research question is the process of intentional inhibition (Brass & Haggard, 2007, 2008; Filevich, Kühn, & Haggard, 2012; Haggard, 2008; Kühn, Haggard, & Brass, 2009; Walsh, Kühn, Brass, Wenke, & Haggard, 2010; Zhang, Hughes, & Rowe, 2012). This research has provided detailed information on how difficult it is to inhibit an action intentionally when there is no awareness. According to neurological evidence it is possible to trigger a (positive) potential selection of an action mainly by the unconscious processes, but it is essential to have conscious elements to enable generation of a (negative) predicted impact that leads to an action inhibition (Filevich, Kühn, & Haggard, 2012). Similar to intentional inhibition, emotional processes are also useful to explore in this question (McRae, Misra, Prasad, Pereira, & Gross, 2012; Ochsner et al., 2009; Weinberg, Ferri, & Hajcak, 2013). In certain situations having a prior awareness is useful to
strengthen or weaken an action based on its predicted effect. In contrast, in certain situations (e.g., flight or fight situations) awareness may not contribute to change or improve the choice for action execution. Therefore, models for cognitive processes were designed considering both claims: awareness has an effect on action execution and it has not. Chapters 2, 3, 4, and 9 provide detailed information about this.

14.1.2. What are the roles of emotions, ownership and cognitive control in human action selection?

In addition to the role of awareness on human action selection some other cognitive concepts also play key roles. Among them, a major role is played by ownership (Treur, 2012), emotions (McRae, Misra, Prasad, Pereira, & Gross, 2012; Ochsner et al., 2009; Weinberg, Ferri, & Hajcak, 2013) and cognitive control (Haggard, 2008; Miller & Cohen, 2001; Miller, 2000; Lavie & Tsal, 1994; Lavie, Hirst, de Fockert, & Viding, 2004). When exploring the research question what is the role of action ownership in action selection, a main question is to differentiate in how far a person attributes an action to him or herself, or to another person (Treur, 2012). The information about another person’s behaviour influences your self-evaluation and vice versa (Chaminade, Marchant, Kilner, & Frith, 2012; Decety & Sommerville, 2003). This separation of self and other is contributing to the ability to recognise ownership of actions and in particular your own actions. Research has provided evidence that action prediction (based on sensory information) plays a crucial role in action execution with ownership; when there are problems with action prediction, that often leads to abnormal states of ownership of that action. Chapters 2, 4, and 9 provide information about the role of ownership in action selection.

Chapters 5 and 9 provide answers for what can be learnt from cognitive/affective/behavioural sciences on how perception, attention, intention, emotion and awareness contribute to action formation and cognitive control. The brain’s circuits for cognitive control consist of loops rather than linear chains (Haggard, 2008). The prefrontal cortex (PFC) plays an important role in top-down driven cognitive control, as a temporal integrator (Miller & Cohen, 2001). The higher order interconnectivity of the PFC with other cortical, and subcortical areas has been interpreted as indicating a process that generates and maintains information when sensory inputs are weak, ambiguous, rapidly changing, novel and/or multiple options exist (Miller, 2000; Miller & Cohen, 2001). Load theory of attention and cognitive control (Lavie & Tsal, 1994) provides detailed information about early versus late selection schemes together with the support of perception. A person under high perceptual load is unable to shift his or her perception to other salient features in the environment. Instead, when the perceptual load is lower, the agent is capable of perceiving information in parallel (Lavie, 2005, 2006; Lavie,
Hirst, de Fockert, & Viding, 2004). Detailed literature on these cognitive processes is included in Chapters 5 and 9 (some other chapters also contribute to this). Chapter 4 addresses the cognitive/affective/behavioural sciences literature on emotion generation processes, related to action selection.

How to model the interplay between conscious vs. unconscious processes in human action formation is also explored. The process of human action formation is a complex system that includes influences of both conscious and unconscious processes. Chapters 4 and 5 provide specific cognitive models (both are sharing the same basis, but Chapter 4 focuses on emotional awareness whereas Chapter 5 focuses on cognitive control related to action formation) which include the interplay of conscious and unconscious processes. Action formation emerges more from unconscious processes and these include bottom-up processes (Katsuki & Constantinidis, 2014; McRae, Misra, Prasad, Pereira, & Gross, 2012; Ochsner et al., 2009; Pessoa, 2010; Pourtois, G., Schettino, A., & Vuilleumier, P. (2013); Sheppes & Gross, 2011; Weinberg, Ferri, & Hajcak, 2013). Having bottom-up processes, these will pass information to higher order cognitive states and enable top-down processes (Baluch & Itti, 2011; Engel, Fries, & Singer, 2001; Haggard, 2008; Katsuki & Constantinidis, 2014; Kiefer, 2007; Miller, 2000; Miller & Cohen, 2001; Ochsner et al., 2009; Pourtois, G., Schettino, A., & Vuilleumier, P. (2013); Sheppes & Gross, 2011; Tallon-Baudry, 2012; Weinberg, Ferri, & Hajcak, 2013). This process is used as the core for action formation, and this is strengthened by a number of other supportive processes.

14.1.3. **How to design cognitive models for joint decision making on action selection and what is the role of cognitive metaphors in such models?**

Making decisions together with others is an essential part of human life, especially in social and professional context. Similarly, understanding new things based on what already has been understood is also essential, which is referred as analogy making. There are analogy making models (Kokinov, 1994; Kokinov, Grinberg, Petkov, & Kiryazov, 2008; Hummel & Holyoak, 1996; Hummel & Holyoak, 1997) inspired by neuro-cognitive theories; a new model proposed for this purpose is mainly inspired by the Theravada Buddhist literature. In Theravada Buddhism, an interesting process has been explained that can be used to explain the process of analogy making together with information of cognitive states. The Theory of Five Aggregates (Thera, 1972, 2008) explained in Theravada Buddhism was used for this to model the analogy making process. The Five Aggregates comprise of: Form, Sensation, Perception, Mental-formation and Consciousness (Boisvert, 1995; Thera, 2008). Chapter 6 provides details of what can be learnt from such Buddhist explanations for analogy making and of a cognitive model that was developed to explain the process of analogy making. This model was validated in
the geometry domain. The basics of this analogy making model has also been used in Chapter 7 for matching a metaphor to a situation at hand.

The way we understand many phenomena in our daily lives is metaphorical: one certain mental domain is understood in terms of another mental phenomenon. It is interesting to explore what can be learnt from cognitive metaphors for joint decision making in particular. In Chapter 7, the influence of cognitive metaphors on joint decision making has been examined by combining the concept of a joint decision making process and cognitive metaphor together in a computational social agent model. In this work two core processes were identified as important to model processes of joint decision making: mirror neurons (Iacoboni, 2008; Treur, 2011a, 2011b) and internal simulation (Wolpert, 1997). While mirror neurons prepare a person for an action, internal simulation generates a prediction of the (expected) effects of such a prepared action. Having these two concepts (i.e., mirror neurons and internal simulation) as an integrated process leads to empathic understanding (Damasio, 2005; Goldman, 2006; Hesslow, 2002) as part of joint decision making, thereby integrating the effect of ownership (Moore & Haggard, 2008; Treur, 2012). Having all these processes as underlying processes in joint decision making, it has been found that our metaphorical image of the situation as identified by a form of analogy making has a strongly influence on this (Cardillo, Watson, Schmidt, Kranjec, & Chatterjee, 2012; Carroll & Thomas, 1982; Leary, 1994; Romero & Soria, 2005).

14.1.4. How can the computational models be applied in complex real-world domains?

Computational cognitive modelling has obtained recognition from many research domains. It has been found that nowadays more than 80% of the published articles in theoretical journals in cognitive science are about cognitive modelling (Busemeyer & Diederich, 2010). Computational modelling is considered to be an important pillar for the further development of cognitive science and its related disciplines (Addyman & French, 2012; McClelland, 2009; Shiffrin, 2010). Designing a generic but sufficiently complex cognitive model is always a challenge. This thesis includes cognitive models mainly focusing on aspects related to action selection. Having an interesting cognitive model, it is essential to validate its behaviour with many scenarios by applying it to real-world domains. Two domains were selected for this purpose: the aviation domain and the energy domain. Both domains involve both humans and technical equipment: they involve socio-technical systems. Situation awareness plays an important role in some areas of the aviation domain and it is a challenge to contribute cognitive models for this purpose (Endsley, 1988). A cognitive model was specialized to simulate situation awareness in the aviation domain. To validate the model, 4 scenarios were selected, of which 3
were taken from examples of the Airbus company. The simulation results provide an acceptable level of explanation thus providing confidence to use the model in more complex simulations (see Chapter 9).

In addition to the aviation domain, the energy domain was selected to apply cognitive models. As the scope of the energy domain is quite different from the rest of the work in this thesis, some separate research work was conducted to build some knowledge on modelling energy related physical processes, especially concerning air to water heat pumps (Chapters 10, 11, and 12). This additional work provided good insight to combine energy models together with cognitive models to provide more realistic simulation results. Domestic energy management is the specific real world domain considered. How to improve the analysis of domestic energy management for heating through dynamic computational models is a well-known question in the field (European Commission, Joint Research Centre, & European Technology Platform on Renewable Heating and Cooling (RHC-Platform), 2011).

Due to the cost and complexity in conducting practical experiments it is beneficial to design dynamic computational models that cover the behaviour of dynamic changes of environmental factors together with the behaviour of relevant energy generating systems (e.g., a heat pump). For this purpose the behaviour of air to water heat pumps (Aste, Adhikari, & Manfren, 2013) together with the behaviour of indoor and outdoor temperature behaviour was modelled as a mathematical dynamical system. Chapter 13 presents research for a cognitive simulation driven domestic heating energy management system using an air to water heat pump. It has integrated the earlier developed mathematical model for this heating system’s performance, and a computational dynamic cognitive model for the human’s behaviour which was developed based on evidence from Cognitive Neuroscience. Through this it addresses the importance of the choices of human behaviour on energy usage with a realistic setup through an alternative approach: a simulation-based analysis. As a result, it gives an overall idea on possible impacts and provides additional information on how to motivate persons or correct their biased perceptions in a methodological approach to support lifestyle and lifestyle change in relation to energy management.

A general question in using computational dynamic cognitive models is how to estimate parameters in them, especially given that usually only incomplete empirical features or patterns over time are available. Due to the level of abstraction often used in cognitive models, and limitations in brain imaging techniques it is a nontrivial task to obtain high quality empirical data to validate designed cognitive models. In most of the empirical research, outputs are in the form of incomplete data concerning features or patterns over time. This situation leads to many difficulties to estimate proper values for parameters of the model. Accuracy and correctness of a cognitive model mainly depend on both the chosen abstraction
assumptions and the values of parameters. Chapter 8 presents a solution for such cognitive models using an improved Particle Swarm Optimization (PSO) (Kennedy & Eberhart, 1995; Palupi Rini, Mariyam Shamsuddin, & Sophiyati Yuhaniz, 2011; Poli, Kennedy, & Blackwell, 2007) algorithm combined with Constraint Satisfaction (CS) (Kumar, 1992; Tsang, 1996). In this approach, a parameter estimation problem for a given dynamic computational model is represented as a constraint satisfaction problem and an improved PSO algorithm is used to search for a possible solution for parameter values together with identified partial temporal and static features and patterns.

14.2. Relations between models developed over time

This thesis explores much literature from cognitive/affective/behavioural (neuro)sciences to understand various interesting cognitive processes. Having interesting literature together with sometimes contradictory evidence makes the modelling process more complex. The scope of the work is limited to human action selection processes. This research started with the question how we actually select our actions and why we select a particular action in a given situation instead of something else (especially in complex situations). In Chapter 2 two hypotheses were integrated:

1. Awareness of action selection is not directly causing execution of the action, but comes afterwards (Baumeister, Masicampo, & Vohs, 2011; D’Ostilio & Garraux, 2012; Haynes, 2011; Libet, Gleason, Wright, & Pearl, 1983; Wegner, 2002), and

2. Predictive and inferential processes related to action preparation and execution may contribute to the conscious awareness of the action (Desantis, Hughes, & Waszak, 2012; Haggard, Clark, & Kalogeras, 2002; Moore, Lagnado, Deal, & Haggard, 2009; Moore & Obhi, 2012; Walsh & Haggard, 2013)

These hypotheses include both conscious and unconscious explanations for both action awareness and ownership and provide relations among interactions between processes through cognitive states. Together with some other literature, these hypotheses led to the design of a cognitive model for action awareness focusing on action preparation and performance. Furthermore, it was made sure that the designed model is very generic and able to simulate many interesting scenarios. This model is the point of departure for the rest of the work.

More cognitive processes were explored through cognitive/affective/behavioural (neuro)sciences literature. The next process considered was intentional inhibition and the findings of that process were nicely fitting into the previously developed model while providing more confidence on the generic nature of the initial model and its strength in extendibility for new processes. The main contributions of this
work (relative to the previous model) were the incorporation of the interplay between a positive potential selection of an action and the negative impacts of the same action, together with performative desires and constitutive desires. This resulted in Chapter 3. Furthermore, this work was extended by incorporating emotional awareness and emotional generation related processes. This piece of work is an extension from the previous work, by incorporating more information in the model. Incorporating the bottom-up and top-down processes were the main contributions. New states, emotional awareness, perception, and attention were introduced in this model, which has led to Chapter 4. The model in Chapter 4 includes many details for action selection and, furthermore, it incorporates cognitive control as well. The process behind cognitive control is a further extension to bottom-up and top-down process incorporating conscious intention and a few other mechanisms (bottom-up and top-down attention, suppression and exciting mechanisms). Chapter 5 presents the work addressing this.

This model presented in Chapter 5 is an advanced model that has the ability to explain many situations related to human action selection processes. Therefore, this model was selected to be applied in an aviation application domain to simulate some selected scenarios; this model was further extended to incorporate processes behind bias perception. These simulation results provide interesting information and were able to explain the three levels that Endsley put forward for SA (the model is slightly different from what Endsley proposed but still is able to simulate the same situations with the same expected behaviour). This work is presented in Chapter 9. Furthermore the same model developed in Chapter 5 was used in another application domain: the domestic energy domain. In general, in the literature, cognitive modelling and energy management are not strongly connected, although human behaviour is a key factor behind heavy domestic energy usage. Therefore, this domain was selected to conduct simulations that provide inputs to explore further research in this integrative line of research. This research integrated the earlier developed mathematical model for a heating system’s performance, and the computational dynamic cognitive model presented in Chapter 5. Through this, it addressed the importance of the choices of human behaviour on energy usage within a realistic setup. This research shows the usefulness of cognitive models in different domains as a simulation component.

In addition, two models were developed for analogy making and joint decision making. The analogy making model was inspired by Theravada Buddhist literature and this shows the possibility of adapting different types of literature into the cognitive modelling. The joint decision making model also uses the basic unconscious processes proposed in Chapter 2 and further extends them by incorporating cognitive metaphors. This work provided another exploration of extendibility of the models in the sense of how interaction among cognitive models
for multiple agents can be addressed, which is different from the other models: in all other chapters, behaviour of a single agent is considered.

14.3. The applicability of the research in this thesis

Computational cognitive modelling is a form of multi-disciplinary research that has many benefits in various domains. Therefore, computational modelling is considered to be important for the development of cognitive science and its related disciplines (Addyman & French, 2012; McClelland, 2009; Shiffrin, 2010). The development in brain imaging and recording techniques allows the research community to uplift the understanding of brain processes and mechanisms (Posner & Raichle, 1999; Raichle, 2003; Kwong et al., 1992). Therefore, information available to understand and express complex cognitive processes is becoming available in more and more detail. Nevertheless, most research in cognitive neuroscience focuses on relatively small but highly important questions using increasingly detailed data (Bassett & Gazzaniga, 2011). It is not straightforward to find detailed explanations of more complex phenomena such as action awareness as an integrated, coherent system. A process of one selected phenomenon is always affected by many others and vice versa, due to the high order of coupling and associations in the human brain. Therefore, there is room in brain-related research areas to design and model compound systems by integrating various types of information collected from various neuro-cognitive research findings and theories.

This type of works is not completely new and, for example, even in 70s also there is literature about this topic. For example Kahneman, & Tversky (1973) presented two different ways the brain forms thoughts (System 1: fast, automatic, frequent, emotional, stereotypic, subconscious; System 2: slow, effortful, infrequent, logical, calculating, conscious). In addition, there are some other models, which are mainly component based. For example, Norman, & Shallice (1986) is an example for this. Nevertheless, models presented in this thesis have taken a different perspective and mainly used relatively recent neurocognitive literature which are achieved through and justified by fMRI data.

The relation between actual human cognition and empirical evidence collected through brain imaging and recording techniques has to bridge two different levels of abstraction. Therefore it is important to explore ways to handle this gap and to transform interesting and useful information into a different form that can be easily understandable and expressible. Dynamic modelling approaches provide insight for this and having a model, designed at a cognitive level but based on underlying neurological knowledge, that can explain cognitive phenomena and situations will be useful in multidisciplinary context. Having developed interesting models from knowledge gathered from cognitive, behavioural and affective sciences, there may be a concern whether these models bring any further in understanding these
processes for cognitive, behavioural and affective science researchers. For example, this type of models can be used as part of a workbench at a more abstract and global level, in order to scrutinize different hypotheses and to explore the mechanisms of different processes integrated as a coherent system occurring in experiments in cognitive, behavioural, and affective sciences. In many application domains (especially safety-critical domains such as the aviation domain), it is not possible to conduct live experiments due to the risks attached and high costs. For such applications complex simulations is a useful alternative solution. Furthermore, even for some complex real-world experiments it also may be a best practice to first conduct a simulation-based experiment and then based on the information and knowledge gathered through it, to design and conduct the actual experiment to save the cost, time, and risk. In these ways simulation-based experiments can play a vital role in today’s context.

Having noted the importance of simulation-based experiments it is a research question to explore what are the main ingredients to design complex simulations (which is not in the scope of this thesis). Among them, one key element is the knowledge that should integrate into the simulation about the domain (for example: how to land an aircraft in a very bad weather condition). For this purpose different knowledge representation techniques can be used, and one interesting method is model-based knowledge representation. Therefore, for better simulations it is necessary to have good models that explain the behaviour of the process. Most phenomena in (human or non-human) nature are quite complex and form interesting research areas for many scientists. Model-based methods have become a major approach for many disciplines to study the complexity in their fields. A complex phenomenon can be modelled in terms of the dynamics and interactions of its states and components more scientifically, with cause and effect relations (causality) as a background. A collection (or network) of such assumed causal relations can be compiled in the form of a dynamical model, which is more open to analysis than the real world phenomenon. The main advantage of such a model is that, while it covers the inherent complexity analogous to the real phenomenon, yet it is simple enough to understand how it works and its emergent or salient features (Maria, 1997). Therefore, a dynamical model which is aligned with a real world phenomenon is a good basis to analyse, assess, or predict the processes of the real world phenomenon. Especially for brain-related cognitive process this model-based approach is a working methodology.

In many complex real world problems human and in particular cognitive factors are integrated into the problem. Due to the complexity and the many cyclic loops and higher order couplings in cognitive processes a dynamical systems modelling perspective is needed to represent knowledge of a cognitive process as a model. In principle, in cognitive modelling, a phenomenon of human cognition (or behaviour)
is represented as (computational) mathematical models to understand the causality thereby using adjustable parameters that can be estimated on the basis of neuro-cognitive findings or theories (Addyman & French, 2012; Busemeyer & Diederich, 2010; Lewandowsky, 2011; McClelland, 2009; Shiffrin, 2010). These cognitive models provide more flexibility to deal with the complexity involved in human brain processes and contribute to the aggregation of interesting findings from brain research. The relation between actual human cognition and empirical evidence collected through brain imaging and recording techniques bridges the two different levels of abstraction. Therefore it is important to explore alternatives to handle this gap and to translate interesting and useful information into a different form that can be easily understandable and expressible. Dynamic modelling approaches provide insight for this and having a model that can explain many cognitive phenomena and situations will be useful in multidisciplinary context.

Starting with a basic cognitive model (see Chapter 2) and extending it to a very detailed and to a broader scale is one main achievement of this work. This creates room to apply this model in many domains where human action selection is needed to be integrated. Especially for complex simulations this type of model is very useful. Also another important element in this research is parameter estimation. Every cognitive model’s behaviour is scrutinised with a few scenarios (for some with even 8 scenarios) and it is made sure to use a generic unique parameter value set for each model. Each scenario of that model is achieved with a very minimal change to the selected parameters. Therefore the model was not fined-tuned to a very specific case but always it was tried to find a generic unique parameter value set. Furthermore, in all these models knowledge of behavioural patterns were not included in the model itself. The knowledge of these patterns was used to validate the generated behaviour of the model through simulations using different values of the parameters. This approach enables justification that a model and its parameters indeed provide the needed insight into a neurological or cognitive process. By designing a computational cognitive model based on the latest neuro-cognitive evidence, it can provide valuable simulation results for different scenarios based on the unique parameter value set; this provides more reliability and confidence on the designed model’s correctness. This feature is very important in complex simulations. In complex simulations it is required to include the effects of many situations and having this feature provides more configurability even at runtime.

This research includes two methods for parameter estimation: analytical driven and an improved PSO algorithm. These two methods are beneficial for domains where only limited empirical data available and features of patterns expected are in non-quantitative and discrete forms. This is one of the main difficulties in dynamic cognitive modelling domain. The quality of parameter estimation achieved by these two methods reflected in simulation results obtained for each model, and these
results are impressive. The applicability of these two methods is not restricted to the
dynamic cognitive modelling domain or related domains. In principle these
techniques (especially the improved PSO algorithm) can be used for parameter
estimation in many applications. The methodology is not restricted to dynamic
cognitive models. Furthermore, the PSO algorithm is especially useful for
applications with multiple interesting near optimal solutions.

Having a detailed compound cognitive model, it can also be used in a
workbench to scrutinize various hypotheses in cognitive neuroscience research
(McClelland, 2009) and this can in turn contribute to improving the quality of
empirical research, especially through the insights that can be collected through
interesting simulations. There are many implications of various hypotheses, theories
or findings, and it would be beneficial if there was a mechanism that can be used to
scrutinize these ideas by using this as a workbench at a much more abstract and
global level (McClelland, 2009). Additionally, the human brain and its phenomena
concern immeasurably complex systems and processes that involve uncountable
many factors that make experiments not always coherent with reality. Nevertheless,
having computational models enables to uplift the progress of understanding these
processes in a broader level as a multidisciplinary approach (Shiffrin, 2010).
Therefore, this type of research can also strongly contribute to the development of
cognitive neuroscience research.

14.4. Future work

The research presented in this thesis includes many achievements, and
contributes to the development of dynamic computational cognitive modelling.
Nevertheless, it is not free from limitations and leaves room for future work. The
cognitive models presented in this thesis were mainly validated through
simulations, although there is room for more validation and especially for empirical
data driven validation. Dynamic cognitive model validation is always a challenge
due to the limitations in measuring the emergence of internal cognitive states and
complexities in human brain processes. Nevertheless, more advanced experiments
collaborating with neuro-cognitive scholars may provide some new data to
scrutinize the workings of these models further. One possible approach for this will
be to use these models for cognitive disorders (maybe with some necessary
extensions or adaptations of the models). There is much research addressing various
cognitive disorders (e.g., Schizophrenia, Alzheimer, and Bipolar Disorder) that
often generates empirical data on internal states of patients. Therefore, interaction
with these areas provides benefits for cognitive modellers as they can get more
empirical data on internal states to use for model validations. Cognitive researchers
can get generic but more detailed models to be used as tools to experiment and to
design new experiments through intuitions they gathered through model based
simulations. The main advantage of a model is that, while to a certain extent it has the inherent complexity analogous to the real phenomenon, yet it is simple enough to understand how it works and its emergent/salient features (Maria, 1997). Therefore, a dynamical model which is aligned with a real world phenomenon is a good basis to analyse, assess, or predict the processes of the real world phenomenon.

In addition to validation by applying these models in more complex domains (where human cognition is necessary), also the design of more complex simulations is another area for future work. The models include the necessary elements for such work although there are many specific factors and knowledge required for this (e.g., specific domain knowledge, and supportive simulation platforms/engines). There are many domains for which it is impossible to conduct sophisticated real world experiments for some research questions (e.g., the aviation domain). Computational simulations strongly contribute to overcome such limitations and provide realistic information to take necessary actions and to reduce the risk associated with those problems. As discussed earlier, such domains lack realistic methods to include human cognition related factors. The type of dynamic computational models covered here has many benefits for such domains. Therefore, more realistic and advanced complex simulations for safety critical socio-technical systems are a promising area for future work.

The current cognitive models presented in this thesis have covered many areas and aspects, including action selection, intentional inhibition, emotional generation, cognitive control, analogy making and cognitive metaphor, and joint decision making. Another interesting area to consider is human-like learning. Learning may contribute to increasing the spectrum of usage of the current cognitive models. Adaptability is a key factor in human cognition and it is interesting to understand the processes behind this in relation to human-like cognitive learning mechanisms. Furthermore, studying how humans gain knowledge through learning helps to contribute to the development of AI. Having cognitive models that include cognitive learning processes may lead to the design and development of more realistic human-like or human-aware AI applications. In terms of the scope of the current thesis that includes action selection and awareness, human error awareness related literature may be one of the points of departure to get into learning in dynamic cognitive models.

Because cognitive modelling is a multi-disciplinary research domain, its development should not be restricted to computer science researchers developing cognitive models. With the rapid developments in human cognition and neuroscience related research, more and more information on cognitive and neurological processes will be available and for many researchers it may be interesting to design and develop cognitive models to uplift both their
understanding and to be used as a workbench for their work. A dynamic cognitive model of the type considered here is based on a temporal-causal network which includes states and links between them, and the processes emerge based on the interactions between these states through the links. It is not difficult to develop a dedicated software system which is able to provide necessary tools to draw a cognitive model. Having drawn a cognitive model, the next phase is the compilation into an executable form. For this particular requirement, and to achieve a homogenous platform for computational cognitive science, it has advantages to develop a high level language which is able to express the knowledge for a cognitive model. A generic translation engine can be developed which is able to transform a conceptually drawn cognitive model into a computer understandable numerical representation. Furthermore, this type of language can be useful to automate other things related to computational cognitive models as well. For example, parameter estimation can be addressed: having a high-level language to express cognitive models easily leads to express designed parameter estimation problem as a constraint satisfaction problem (modellers or users can express expected patterns or features as constraints through this language) and then be able to apply parameter estimation methods directly.

Cognitive architectures are also an influential and useful line of research for the development of a theoretical, computational cognitive science discipline. Another area of future work is to compare the behaviour of the cognitive models developed in this thesis with that of other cognitive architectures (e.g., ACT-R, Soar, and Clarion). Most cognitive architectures are based on well-established cognitive hypotheses and theories, but are highly abstract and condensed to only one specific hypothesis or theory. In contrast, the models in this thesis are based on combining many (relatively new) cognitive/neurological findings and hypotheses. Therefore, having such benchmarks will provide more detail to understand the usefulness of each approach together with directions to further development in cognitive modelling. Furthermore, with the proposed high level language for computational cognitive models, there is another possibility of integrating some aspects of cognitive architectures and the type of models presented in this thesis. This type of research has potential to uplift the current level of the state of art in theoretical, computational cognitive modelling.
References


http://doi.org/10.1016/S1364-6613(02)01913-7


http://doi.org/10.1177/1073858413514136

http://dx.doi.org/10.1037/h0034747

http://doi.org/10.1109/ICNN.1995.488968


http://doi.org/10.1002/hbm.20711


http://doi.org/10.1073/pnas.89.12.5675


http://doi.org/10.1037/0096-3445.133.3.339
http://doi.org/10.3758/BF03213897


http://doi.org/10.1093/brain/106.3.623


http://doi.org/10.1111/j.1756-8765.2008.01003.x

http://doi.org/10.1016/S0893-6080(00)00059-9

http://doi.org/10.1093/scan/nsq103

http://doi.org/10.1093/brain/121.6.1013

http://doi.org/10.1038/35036228


http://doi.org/10.1016/j.concog.2006.12.004

http://doi.org/10.1016/j.cognition.2008.11.006


Thera, P. (1972). The Psychological Aspect of Buddhism. BPS.


සාරාංශය

381

සාරාංශය

උගත අන්වතියේ තන්ත්‍රිය පිලිබද මඟින් අතර මෙම මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟි

සංශේෂණයේ මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟි

සංශේෂණයේ මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟි

සංශේෂණයේ මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟි

සංශේෂණයේ මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟි

සංශේෂණයේ මඟින් මඟින් මඟින් මඟින් මඟි

සංශේෂණයේ මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟි

සංශේදය

උගත අන්වතියේ තන්ත්‍රිය පිලිබද මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟි

උගත අන්වතියේ තන්ත්‍රිය පිලිබද මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟින් මඟි

උගත අන්වතියේ තන්ත්‍රිය පිලිබද මඟි

උගත අන්වතියේ තන්ත්‍රිය පිලිබද මඟි

උගත අන්වතියේ තන්ත්‍රීය පිලිබද මඟි

උගත අන්වතියේ තන්ත්‍රීය පිලිබද මඟි

උගත අන්වතියේ තන්ත්‍රීය පිලිබද මඟි

උගත අන්වතියේ තන්ත්‍රීය පිලිබද මඟි
ඉහලට බලපෑම, සංශාරය සහ බලශ්කම්, කිහිපයක් නිර්මාණ කර ගනී, වෙනස්වේන් ලෙස සැතන්නවේන්වත්තා බෝධ ලෙස සිතියන්න පිලිබඳ අභ්ංසක මහ පාත්‍ර විශේෂීය අශ්ටක පිළිබද නම්. මහ නිර්මාණය ගැනීම සැම්බරින් ලෙස ලෙස දොරිතුම්, කීව විනිස්කලිත වන්නේ බෝධ අරමුණ පිළිබද අන්තර්ගතික යුතු නිරූපිණිතයක් ආදිවාසය මහ ද කොට ගැනීම් අත් තොරතුරු දක්වා ඉන්නේ පෙක් මෙන්කා පද්ධතික කරුණි.

සිරිම මෙන්නේ ආරම්භ කරුණි මෙන්නේ බෝධ ක්ෂේත්‍ර ආදිවාස පරිදි පවතින දෙළේ ක්ෂේත්‍රක මහ ලෙස දොරිතුම් දක්වා ඉන්නේ ආදිවාසයක් මහ ක්ෂේත්‍ර සිතියන්නේ දෙන්වත්තා තුන්තෙන්හ බෝධ. පවතින මහ ක්ෂේත්‍ර සිතියන්නේ මහ ප්‍රතිඵලිත සිතියන්නේ දෙන්වත්තා තුන්තෙන්හ බෝධ ක්ෂේත්‍ර, අම සිතියන්නේ බෝධ පවතින මහ ප්‍රතිඵලිත සිතියන්නේ දෙන්වත්තා තුන්තෙන්හ බෝධ.
SIKS dissertation series

2009

2009-01  Rasa Jurgelenaite (RUN)
             *Symmetric Causal Independence Models*

2009-02  Willem Robert van Hage (VU)
             *Evaluating Ontology-Alignment Techniques*

2009-03  Hans Stol (UvT)
             *A Framework for Evidence-based Policy Making Using IT*

2009-04  Josephine Nabukunya (RUN)
             *Improving the Quality of Organisational Policy Making using Collaboration Engineering*

2009-05  Sietse Overbeek (RUN)
             *Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality*

2009-06  Muhammad Subianto (UU)
             *Understanding Classification*

2009-07  Ronald Poppe (UT)
             *Discriminative Vision-Based Recovery and Recognition of Human Motion*

2009-08  Volker Nannen (VU)
             *Evolutionary Agent-Based Policy Analysis in Dynamic Environments*

2009-09  Benjamin Kanagwa (RUN)
             *Design, Discovery and Construction of Service-oriented Systems*

2009-10  Jan Wielemaker (UVA)
             *Logic programming for knowledge-intensive interactive applications*

2009-11  Alexander Boer (UVA)
             *Legal Theory, Sources of Law & the Semantic Web*

2009-12  Peter Massuthe (TUE, Humboldt-Universitaet zu Berlin)
             *Operating Guidelines for Services*

2009-13  Steven de Jong (UM)
             *Fairness in Multi-Agent Systems*

2009-14  Maksym Korotkiy (VU)
             *From ontology-enabled services to service-enabled ontologies (making ontologies work in eScience with ONTO-SOA)*

2009-15  Rinke Hoekstra (UVA)
             *Ontology Representation - Design Patterns and Ontologies that Make Sense*

2009-16  Fritz Reul (UvT)
             *New Architectures in Computer Chess*

2009-17  Laurens van der Maaten (UvT)
             *Feature Extraction from Visual Data*

2009-18  Fabian Groffen (CWI)
Armada, An Evolving Database System

2009-19  Valentin Robu (CWI)
Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets

2009-20  Bob van der Vecht (UU)
Adjustable Autonomy: Controlling Influences on Decision Making

2009-21  Stijn Vanderlooy (UM)
Ranking and Reliable Classification

2009-22  Pavel Serdyukov (UT)
Search For Expertise: Going beyond direct evidence

2009-23  Peter Hofgesang (VU)
Modelling Web Usage in a Changing Environment

2009-24  Annerieke Heuvelink (VUA)
Cognitive Models for Training Simulations

2009-25  Alex van Ballegooij (CWI)
"RAM: Array Database Management through Relational Mapping"

2009-26  Fernando Koch (UU)
An Agent-Based Model for the Development of Intelligent Mobile Services

2009-27  Christian Glahn (OU)
Contextual Support of social Engagement and Reflection on the Web

2009-28  Sander Evers (UT)
Sensor Data Management with Probabilistic Models

2009-29  Stanislav Pokraev (UT)
Model-Driven Semantic Integration of Service-Oriented Applications

2009-30  Marcin Zukowski (CWI)
Balancing vectorized query execution with bandwidth-optimized storage

2009-31  Sofiya Katrenko (UVA)
A Closer Look at Learning Relations from Text

2009-32  Rik Farenhorst (VU) and Remco de Boer (VU)
Architectural Knowledge Management: Supporting Architects and Auditors

2009-33  Khiet Truong (UT)
How Does Real Affect Affect Recognition In Speech?

2009-34  Inge van de Weerd (UU)
Advancing in Software Product Management: An Incremental Method Engineering Approach

2009-35  Wouter Koelewijn (UL)
Privacy en Politiegegevens; Over geautomatiseerde normatieve informatie-uitwisseling

2009-36  Marco Kalz (OUN)
Placement Support for Learners in Learning Networks

2009-37  Hendrik Drachsler (OUN)
Navigation Support for Learners in Informal Learning Networks
2009-38 Riina Vuorikari (OU)
Tags and self-organisation: a metadata ecology for learning resources in a multilingual context

2009-39 Christian Stahl (TUE, Humboldt-Universitaet zu Berlin)
Service Substitution -- A Behavioral Approach Based on Petri Nets

2009-40 Stephan Raaijmakers (UvT)
Multinomial Language Learning: Investigations into the Geometry of Language

2009-41 Igor Berezhnyy (UvT)
Digital Analysis of Paintings

2009-42 Toine Bogers (UvT)
Recommender Systems for Social Bookmarking

2009-43 Virginia Nunes Leal Franqueira (UT)
Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients

2009-44 Roberto Santana Tapia (UT)
Assessing Business-IT Alignment in Networked Organizations

2009-45 Jilles Vreeken (UU)
Making Pattern Mining Useful

2009-46 Loredana Afanasiev (UvA)
Querying XML: Benchmarks and Recursion

2010

2010-01 Matthijs van Leeuwen (UU)
Patterns that Matter

2010-02 Ingo Wassink (UT)
Work flows in Life Science

2010-03 Joost Geurts (CWI)
A Document Engineering Model and Processing Framework for Multimedia documents

2010-04 Olga Kulyk (UT)
Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments

2010-05 Claudia Hauff (UT)
Predicting the Effectiveness of Queries and Retrieval Systems

2010-06 Sander Bakkes (UvT)
Rapid Adaptation of Video Game AI

2010-07 Wim Fikkert (UT)
Gesture interaction at a Distance

2010-08 Krzysztof Siewicz (UL)
Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments

2010-09 Hugo Kielman (UL)
2010-10  Rebecca Ong (UL)  
*A Politieke gegevensverwerking en Privacy, Naar een effectieve waarborging*

2010-11  Adriaan Ter Mors (TUD)  
*Mobile Communication and Protection of Children*

2010-12  Susan van den Braak (UU)  
*Sensemaking software for crime analysis*

2010-13  Gianluigi Folino (RUN)  
*High Performance Data Mining using Bio-inspired techniques*

2010-14  Sander van Splunter (VU)  
*Automated Web Service Reconfiguration*

2010-15  Lianne Bodenstaff (UT)  
*Managing Dependency Relations in Inter-Organizational Models*

2010-16  Sicco Verwer (TUD)  
*Efficient Identification of Timed Automata, theory and practice*

2010-17  Spyros Kotoulas (VU)  
*Scalable Discovery of Networked Resources: Algorithms, Infrastructure, Applications*

2010-18  Charlotte Gerritsen (VU)  
*Caught in the Act: Investigating Crime by Agent-Based Simulation*

2010-19  Henriette Cramer (UvA)  
*People’s Responses to Autonomous and Adaptive Systems*

2010-20  Ivo Swartjes (UT)  
*Whose Story Is It Anyway? How Improv Informs Agency and Authorship of Emergent Narrative*

2010-21  Harold van Heerde (UT)  
*Privacy-aware data management by means of data degradation*

2010-22  Michiel Hildebrand (CWI)  
*End-user Support for Access to Heterogeneous Linked Data*

2010-23  Bas Steunebrink (UU)  
*The Logical Structure of Emotions*

2010-24  Dmytro Tykhonov  
*Designing Generic and Efficient Negotiation Strategies*

2010-25  Zulfqar Ali Memon (VU)  
*Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective*

2010-26  Ying Zhang (CWI)  
*XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines*

2010-27  Marten Voulon (UL)  
*Automatisch contracteren*

2010-28  Arne Koopman (UU)  
*Characteristic Relational Patterns*

2010-29  Stratos Idreos(CWI)  
*Database Cracking: Towards Auto-tuning Database Kernels*
<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-30</td>
<td>Accessing Natural History - Discoveries in data cleaning, structuring, and retrieval</td>
<td>Marieke van Erp (UvT)</td>
</tr>
<tr>
<td>2010-31</td>
<td>Ontology Enrichment from Heterogeneous Sources on the Web</td>
<td>Victor de Boer (UVA)</td>
</tr>
<tr>
<td>2010-32</td>
<td>An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems</td>
<td>Marcel Hiel (UvT)</td>
</tr>
<tr>
<td>2010-33</td>
<td>Modeling Representation Uncertainty in Concept-Based Multimedia Retrieval</td>
<td>Robin Aly (UT)</td>
</tr>
<tr>
<td>2010-34</td>
<td>Interaction Design in Service Compositions</td>
<td>Teduh Dirghayu (UT)</td>
</tr>
<tr>
<td>2010-35</td>
<td>Proof of Concept: Concept-based Biomedical Information Retrieval</td>
<td>Dolf Trieschnigg (UT)</td>
</tr>
<tr>
<td>2010-36</td>
<td>Paving the Way for Lifelong Learning: Facilitating competence development through a learning path specification</td>
<td>Jose Janssen (OU)</td>
</tr>
<tr>
<td>2010-37</td>
<td>Correctness of services and their composition</td>
<td>Niels Lohmann (TUE)</td>
</tr>
<tr>
<td>2010-38</td>
<td>From Scenarios to components</td>
<td>Dirk Fahland (TUE)</td>
</tr>
<tr>
<td>2010-39</td>
<td>Integrative modeling of emotions in virtual agents</td>
<td>Ghazanfar Farooq Siddiqui (VU)</td>
</tr>
<tr>
<td>2010-40</td>
<td>Converting and Integrating Vocabularies for the Semantic Web</td>
<td>Mark van Assem (VU)</td>
</tr>
<tr>
<td>2010-41</td>
<td>Monte-Carlo Tree Search</td>
<td>Guillaume Chaslot (UM)</td>
</tr>
<tr>
<td>2010-42</td>
<td>Needs-driven service bundling in a multi-supplier setting - the computational e3-service approach</td>
<td>Sybren de Kinderen (VU)</td>
</tr>
<tr>
<td>2010-43</td>
<td>A Computational Approach to Content-Based Retrieval of Folk Song Melodies</td>
<td>Peter van Kranenburg (UU)</td>
</tr>
<tr>
<td>2010-44</td>
<td>An Approach towards Context-sensitive and User-adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain</td>
<td>Pieter Bellekens (TUE)</td>
</tr>
<tr>
<td>2010-45</td>
<td>A theory and model for the evolution of software services</td>
<td>Vasilios Andrikopoulos (UvT)</td>
</tr>
<tr>
<td>2010-46</td>
<td>e3alignment: Exploring Inter-Organizational Business-ICT Alignment</td>
<td>Vincent Pijpers (VU)</td>
</tr>
<tr>
<td>2010-47</td>
<td>Mining Process Model Variants: Challenges, Techniques, Examples</td>
<td>Chen Li (UT)</td>
</tr>
<tr>
<td>2010-48</td>
<td>Withdrawn</td>
<td></td>
</tr>
<tr>
<td>2010-49</td>
<td>Solving difficult game positions</td>
<td>Jahn-Takeshi Saito (UM)</td>
</tr>
</tbody>
</table>
2010-50  Bouke Huurnink (UVA)
Search in Audiovisual Broadcast Archives

2010-51  Alia Khairia Amin (CWI)
Understanding and supporting information seeking tasks in multiple sources

2010-52  Peter-Paul van Maanen (VU)
Adaptive Support for Human-Computer Teams: Exploring the Use of Cognitive Models of Trust and Attention

2010-53  Edgar Meij (UVA)
Combining Concepts and Language Models for Information Access

2011

2011-01  Botond Cseke (RUN)
Variational Algorithms for Bayesian Inference in Latent Gaussian Models

2011-02  Nick Tinnemeier(UU)
Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language

2011-03  Jan Martijn van der Werf (TUE)
Compositional Design and Verification of Component-Based Information Systems

2011-04  Hado van Hasselt (UU)
Insights in Reinforcement Learning: Formal analysis and empirical evaluation of temporal-difference learning algorithms

2011-05  Base van der Raadt (VU)
Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline.

2011-06  Yiwen Wang (TUE)
Semantically-Enhanced Recommendations in Cultural Heritage

2011-07  Yujia Cao (UT)
Multimodal Information Presentation for High Load Human Computer Interaction

2011-08  Nieske Vergunst (UU)
BDI-based Generation of Robust Task-Oriented Dialogues

2011-09  Tim de Jong (OU)
Contextualised Mobile Media for Learning

2011-10  Bart Bogaert (UvT)
Cloud Content Contention

2011-11  Dhaval Vyas (UT)
Designing for Awareness: An Experience-focused HCI Perspective

2011-12  Carmen Bratosin (TUE)
Grid Architecture for Distributed Process Mining

2011-13  Xiaoyu Mao (UvT)
Airport under Control. Multiagent Scheduling for Airport Ground Handling

2011-14  Milan Lovric (EUR)
Behavioral Finance and Agent-Based Artificial Markets
2011-15 Marijn Koolen (UvA)
*The Meaning of Structure: the Value of Link Evidence for Information Retrieval*

2011-16 Maarten Schadd (UM)
*Selective Search in Games of Different Complexity*

2011-17 Jiyin He (UVA)
*Exploring Topic Structure: Coherence, Diversity and Relatedness*

2011-18 Mark Ponsen (UM)
*Strategic Decision-Making in complex games*

2011-19 Ellen Rusman (OU)
*The Mind ’ s Eye on Personal Profiles*

2011-20 Qing Gu (VU)
*Guiding service-oriented software engineering - A view-based approach*

2011-21 Linda Terlouw (TUD)
*Modularization and Specification of Service-Oriented Systems*

2011-22 Junte Zhang (UVA)
*System Evaluation of Archival Description and Access*

2011-23 Wouter Weerkamp (UVA)
*Finding People and their Utterances in Social Media*

2011-24 Herwin van Welbergen (UT)
*Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior*

2011-25 Syed Waqar ul Qounain Jaffry (VU)
*Analysis and Validation of Models for Trust Dynamics*

2011-26 Matthijs Aart Pontier (VU)
*Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots*

2011-27 Aniel Bhulai (VU)
*Dynamic website optimization through autonomous management of design patterns*

2011-28 Rianne Kapteijn(UVA)
*Effective Focused Retrieval by Exploiting Query Context and Document Structure*

2011-29 Faisal Kamiran (TUE)
*Discrimination-aware Classification*

2011-30 Egon van den Broek (UT)
*Affective Signal Processing (ASP): Unraveling the mystery of emotions*

2011-31 Ludo Waltman (EUR)
*Computational and Game-Theoretic Approaches for Modeling Bounded Rationality*

2011-32 Nees-Jan van Eck (EUR)
*Methodological Advances in Bibliometric Mapping of Science*

2011-33 Tom van der Weide (UU)
*Arguing to Motivate Decisions*

2011-34 Paolo Turrini (UU)
*Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations*
<table>
<thead>
<tr>
<th>Year</th>
<th>Code</th>
<th>Author (Institution)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>35</td>
<td>Maaike Harbers (UU)</td>
<td>Explaining Agent Behavior in Virtual Training</td>
</tr>
<tr>
<td>2011</td>
<td>36</td>
<td>Erik van der Spek (UU)</td>
<td>Experiments in serious game design: a cognitive approach</td>
</tr>
<tr>
<td>2011</td>
<td>37</td>
<td>Adriana Burlutiu (RUN)</td>
<td>Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference</td>
</tr>
<tr>
<td>2011</td>
<td>38</td>
<td>Nyree Lemmens (UM)</td>
<td>Bee-inspired Distributed Optimization</td>
</tr>
<tr>
<td>2011</td>
<td>39</td>
<td>Joost Westra (UU)</td>
<td>Organizing Adaptation using Agents in Serious Games</td>
</tr>
<tr>
<td>2011</td>
<td>40</td>
<td>Viktor Clerc (VU)</td>
<td>Architectural Knowledge Management in Global Software Development</td>
</tr>
<tr>
<td>2011</td>
<td>41</td>
<td>Luan Ibraimi (UT)</td>
<td>Cryptographically Enforced Distributed Data Access Control</td>
</tr>
<tr>
<td>2011</td>
<td>42</td>
<td>Michal Sindlar (UU)</td>
<td>Explaining Behavior through Mental State Attribution</td>
</tr>
<tr>
<td>2011</td>
<td>43</td>
<td>Henk van der Schuur (UU)</td>
<td>Process Improvement through Software Operation Knowledge</td>
</tr>
<tr>
<td>2011</td>
<td>44</td>
<td>Boris Reuderink (UT)</td>
<td>Robust Brain-Computer Interfaces</td>
</tr>
<tr>
<td>2011</td>
<td>45</td>
<td>Herman Stehouwer (UvT)</td>
<td>Statistical Language Models for Alternative Sequence Selection</td>
</tr>
<tr>
<td>2011</td>
<td>46</td>
<td>Beibei Hu (TUD)</td>
<td>Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work</td>
</tr>
<tr>
<td>2011</td>
<td>47</td>
<td>Azizi Bin Ab Aziz (VU)</td>
<td>Exploring Computational Models for Intelligent Support of Persons with Depression</td>
</tr>
<tr>
<td>2011</td>
<td>48</td>
<td>Mark Ter Maat (UT)</td>
<td>Response Selection and Turn-taking for a Sensitive Artificial Listening Agent</td>
</tr>
<tr>
<td>2011</td>
<td>49</td>
<td>Andreea Niculescu (UT)</td>
<td>Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality</td>
</tr>
</tbody>
</table>

**2012**

<table>
<thead>
<tr>
<th>Year</th>
<th>Code</th>
<th>Author (Institution)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>01</td>
<td>Terry Kakeeto (UvT)</td>
<td>Relationship Marketing for SMEs in Uganda</td>
</tr>
<tr>
<td>2012</td>
<td>02</td>
<td>Muhammad Umair (VU)</td>
<td>Adaptivity, emotion, and Rationality in Human and Ambient Agent Models</td>
</tr>
<tr>
<td>2012</td>
<td>03</td>
<td>Adam Vanya (VU)</td>
<td>Supporting Architecture Evolution by Mining Software Repositories</td>
</tr>
<tr>
<td>2012</td>
<td>04</td>
<td>Jurriaan Souer (UU)</td>
<td></td>
</tr>
</tbody>
</table>
Development of Content Management System-based Web Applications
2012-05  Marijn Plomp (UU)

Maturing Interorganisational Information Systems

2012-06  Wolfgang Reinhardt (OU)

Awareness Support for Knowledge Workers in Research Networks

2012-07  Rianne van Lambalgen (VU)

When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions

2012-08  Gerben de Vries (UVA)

Kernel Methods for Vessel Trajectories

2012-09  Ricardo Neisse (UT)

Trust and Privacy Management Support for Context-Aware Service Platforms

2012-10  David Smits (TUE)

Towards a Generic Distributed Adaptive Hypermedia Environment

2012-11  J.C.B. Rantham Prabhakara (TUE)

Process Mining in the Large: Preprocessing, Discovery, and Diagnostics

2012-12  Kees van der Sluijs (TUE)

Model Driven Design and Data Integration in Semantic Web Information Systems

2012-13  Suleman Shahid (UvT)

Fun and Face: Exploring non-verbal expressions of emotion during playful interactions

2012-14  Evgeny Knutov (TUE)

Generic Adaptation Framework for Unifying Adaptive Web-based Systems

2012-15  Natalie van der Wal (VU)

Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes

2012-16  Fiemke Both (VU)

Helping people by understanding them - Ambient Agents supporting task execution and depression treatment

2012-17  Amal Elgammal (UvT)

Towards a Comprehensive Framework for Business Process Compliance

2012-18  Eltjo Poort (VU)

Improving Solution Architecting Practices

2012-19  Helen Schonenberg (TUE)

What's Next? Operational Support for Business Process Execution

2012-20  Ali Bahramisharif (RUN)

Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing

2012-21  Roberto Cornacchia (TUD)

Querying Sparse Matrices for Information Retrieval

2012-22  Thijs Vis (UvT)

Intelligence, politie en veiligheidsdienst: verenigbare grootheden?

2012-23  Christian Muehl (UT)
Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction

2012-24 Laurens van der Werff (UT)
Evaluation of Noisy Transcripts for Spoken Document Retrieval

2012-25 Silja Eckartz (UT)
Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application

2012-26 Emile de Maat (UVA)
Making Sense of Legal Text

2012-27 Hayrettin Gurkok (UT)
Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games

2012-28 Nancy Pascale (UvT)
Engendering Technology Empowering Women

2012-29 Almer Tigelaar (UT)
Peer-to-Peer Information Retrieval

2012-30 Alina Pommerantz (TUD)
Designing Human-Centered Systems for Reflective Decision Making

2012-31 Emily Bagarukayo (RUN)
A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure

2012-32 Wietske Visser (TUD)
Qualitative multi-criteria preference representation and reasoning

2012-33 Rory Sie (OUN)
Coalitions in Cooperation Networks (COCOON)

2012-34 Pavol Jancaura (RUN)
Evolutionary analysis in PPI networks and applications

2012-35 Evert Haasdijk (VU)
Never Too Old To Learn -- On-line Evolution of Controllers in Swarm- and Modular Robotics

2012-36 Denis Ssebugwawo (RUN)
Analysis and Evaluation of Collaborative Modeling Processes

2012-37 Agnes Nakakawa (RUN)
A Collaboration Process for Enterprise Architecture Creation

2012-38 Selmar Smit (VU)
Parameter Tuning and Scientific Testing in Evolutionary Algorithms

2012-39 Hassan Fatemi (UT)
Risk-aware design of value and coordination networks

2012-40 Agus Gunawan (UvT)
Information Access for SMEs in Indonesia

2012-41 Sebastian Kelle (OU)
Game Design Patterns for Learning

2012-42 Dominique Verpoorten (OU)
Reflection Amplifiers in self-regulated Learning

2012-43 Withdrawn

2012-44 Anna Tordai (VU)
On Combining Alignment Techniques

2012-45 Benedikt Kratz (UvT)
A Model and Language for Business-aware Transactions

2012-46 Simon Carter (UVA)
Exploration and Exploitation of Multilingual Data for Statistical Machine Translation

2012-47 Manos Tsagkias (UVA)
Mining Social Media: Tracking Content and Predicting Behavior

2012-48 Jorn Bakker (TUE)
Handling Abrupt Changes in Evolving Time-series Data

2012-49 Michael Kaisers (UM)
Learning against Learning - Evolutionary dynamics of reinforcement learning algorithms in strategic interactions

2012-50 Steven van Kervel (TUD)
Ontology driven Enterprise Information Systems Engineering

2012-51 Jeroen de Jong (TUD)
Heuristics in Dynamic Scheduling; a practical framework with a case study in elevator dispatching

2013

2013-01 Viorel Milea (EUR)
News Analytics for Financial Decision Support

2013-02 Erietta Liarou (CWI)
MonetDB/DataCell: Leveraging the Column-store Database Technology for Efficient and Scalable Stream Processing

2013-03 Szymon Klarman (VU)
Reasoning with Contexts in Description Logics

2013-04 Chetan Yadati(TUD)
Coordinating autonomous planning and scheduling

2013-05 Dulce Pumareja (UT)
Groupware Requirements Evolutions Patterns

2013-06 Romulo Goncalves(CWI)
The Data Cyclotron: Juggling Data and Queries for a Data Warehouse Audience

2013-07 Giel van Lankveld (UvT)
Quantifying Individual Player Differences

2013-08 Robbert-Jan Merk(VU)
Making enemies: cognitive modeling for opponent agents in fighter pilot simulators

2013-09 Fabio Gori (RUN)
Metagenomic Data Analysis: Computational Methods and Applications

2013-10 Jeewanie Jayasinghe Arachchige(UvT)
A Unified Modeling Framework for Service Design
2013-11 Evangelos Pournaras (TUD)
Multi-level Reconfigurable Self-organization in Overlay Services

2013-12 Marian Razavian (VU)
Knowledge-driven Migration to Services

Service Tailoring: User-centric creation of integrated IT-based homecare services to support independent living of elderly

2013-13 Mohammad Safiri (UT)

2013-14 Jafar Tanha (UVA)
Ensemble Approaches to Semi-Supervised Learning Learning

2013-15 Daniel Hennes (UM)
Multiagent Learning - Dynamic Games and Applications

2013-16 Eric Kok (UU)
Exploring the practical benefits of argumentation in multi-agent deliberation

2013-17 Koen Kok (VU)
The PowerMatcher: Smart Coordination for the Smart Electricity Grid

2013-18 Jeroen Janssens (UvT)
Outlier Selection and One-Class Classification

2013-19 Renze Steenhuizen (TUD)
Coordinated Multi-Agent Planning and Scheduling

2013-20 Katja Hofmann (UvA)
Fast and Reliable Online Learning to Rank for Information Retrieval

2013-21 Sander Wubben (UvT)
Text-to-text generation by monolingual machine translation

2013-22 Tom Claassen (RUN)
Causal Discovery and Logic

2013-23 Patricio de Alencar Silva (UvT)
Value Activity Monitoring

2013-24 Haitham Bou Ammar (UM)
Automated Transfer in Reinforcement Learning

2013-25 Agnieszka Anna Latoszek-Berendsen (UM)
Intention-based Decision Support. A new way of representing and implementing clinical guidelines in a Decision Support System

2013-26 Alireza Zarghami (UT)
Architectural Support for Dynamic Homecare Service Provisioning

2013-27 Mohammad Huq (UT)
Inference-based Framework Managing Data Provenance

2013-28 Frans van der Sluis (UT)
When Complexity becomes Interesting: An Inquiry into the Information eXperience

2013-29 Iwan de Kok (UT)
Listening Heads

2013-30 Joyce Nakatumba (TUE)
Resource-Aware Business Process Management: Analysis and Support
2013-31 Dinh Khoa Nguyen (UvT)
Blueprint Model and Language for Engineering Cloud Applications

2013-32 Kamakshi Rajagopal (OUN)
Networking For Learning; The role of Networking in a Lifelong Learner's Professional Development

2013-33 Qi Gao (TUD)
User Modeling and Personalization in the Microblogging Sphere

2013-34 Kien Tjin-Kam-Jet (UT)
Distributed Deep Web Search

2013-35 Abdallah El Ali (UvA)
Minimal Mobile Human Computer Interaction
Promotor: Prof. dr. L. Hardman (CWI/UVA)

2013-36 Than Lam Hoang (TUe)
Pattern Mining in Data Streams

2013-37 Dirk Börner (OUN)
Ambient Learning Displays

2013-38 Eelco den Heijer (VU)
Autonomous Evolutionary Art

2013-39 Joop de Jong (TUD)
A Method for Enterprise Ontology based Design of Enterprise Information Systems

2013-40 Pim Nijssen (UM)
Monte-Carlo Tree Search for Multi-Player Games

2013-41 Jochem Liem (UVA)
Supporting the Conceptual Modelling of Dynamic Systems: A Knowledge Engineering Perspective on Qualitative Reasoning

2013-42 Léon Planken (TUD)
Algorithms for Simple Temporal Reasoning

2013-43 Marc Bron (UVA)
Exploration and Contextualization through Interaction and Concepts

2014

2014-01 Nicola Barile (UU)
Studies in Learning Monotone Models from Data

2014-02 Fiona Tuliyano (RUN)
Combining System Dynamics with a Domain Modeling Method

2014-03 Sergio Raul Duarte Torres (UT)
Information Retrieval for Children: Search Behavior and Solutions

2014-04 Hanna Jochmann-Mannak (UT)
Websites for children: search strategies and interface design - Three studies on children's search performance and evaluation
<table>
<thead>
<tr>
<th>Year</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-05</td>
<td>Jurriaan van Reijsen (UU)</td>
<td>Knowledge Perspectives on Advancing Dynamic Capability</td>
</tr>
<tr>
<td>2014-06</td>
<td>Damian Tamburri (VU)</td>
<td>Supporting Networked Software Development</td>
</tr>
<tr>
<td>2014-07</td>
<td>Arya Adriansyah (TUE)</td>
<td>Aligning Observed and Modeled Behavior</td>
</tr>
<tr>
<td>2014-08</td>
<td>Samur Araujo (TUD)</td>
<td>Data Integration over Distributed and Heterogeneous Data Endpoints</td>
</tr>
<tr>
<td>2014-09</td>
<td>Philip Jackson (UvT)</td>
<td>Toward Human-Level Artificial Intelligence: Representation and Computation of Meaning in Natural Language</td>
</tr>
<tr>
<td>2014-10</td>
<td>Ivan Salvador Razo Zapata (VU)</td>
<td>Service Value Networks</td>
</tr>
<tr>
<td>2014-11</td>
<td>Janneke van der Zwaan (TUD)</td>
<td>An Empathic Virtual Buddy for Social Support</td>
</tr>
<tr>
<td>2014-12</td>
<td>Willem van Willigen (VU)</td>
<td>Look Ma, No Hands: Aspects of Autonomous Vehicle Control</td>
</tr>
<tr>
<td>2014-13</td>
<td>Arlette van Wissen (VU)</td>
<td>Agent-Based Support for Behavior Change: Models and Applications in Health and Safety Domains</td>
</tr>
<tr>
<td>2014-14</td>
<td>Yangyang Shi (TUD)</td>
<td>Language Models With Meta-information</td>
</tr>
<tr>
<td>2014-16</td>
<td>Krystyna Milian (VU)</td>
<td>Supporting trial recruitment and design by automatically interpreting eligibility criteria</td>
</tr>
<tr>
<td>2014-18</td>
<td>Mattijs Ghijsen (UVA)</td>
<td>Methods and Models for the Design and Study of Dynamic Agent Organizations</td>
</tr>
<tr>
<td>2014-19</td>
<td>Vinicius Ramos (TUE)</td>
<td>Adaptive Hypermedia Courses: Qualitative and Quantitative Evaluation and Tool Support</td>
</tr>
<tr>
<td>2014-21</td>
<td>Kassidy Clark (TUD)</td>
<td>Negotiation and Monitoring in Open Environments</td>
</tr>
<tr>
<td>2014-22</td>
<td>Marieke Peeters (UU)</td>
<td>Personalized Educational Games - Developing agent-supported scenario-based training</td>
</tr>
</tbody>
</table>
2014-23 Eleftherios Sidirourgos (UvA/CWI)
Space Efficient Indexes for the Big Data Era

2014-24 Davide Ceolin (VU)
Trusting Semi-structured Web Data

2014-25 Martijn Lappenschaar (RUN)
New network models for the analysis of disease interaction

2014-26 Tim Baarslag (TUD)
What to Bid and When to Stop

2014-27 Rui Jorge Almeida (EUR)
Conditional Density Models Integrating Fuzzy and Probabilistic Representations of Uncertainty

2014-28 Anna Chmielowiec (VU)
Decentralized k-Clique Matching

2014-29 Jaap Kabbedijk (UU)
Variability in Multi-Tenant Enterprise Software

2014-30 Peter de Cock (UvT)
Anticipating Criminal Behaviour

2014-31 Leo van Moergestel (UU)
Agent Technology in Agile Multiparallel Manufacturing and Product Support

2014-32 Naser Ayat (UvA)
On Entity Resolution in Probabilistic Data

2014-33 Tesfa Tegegne (RUN)
Service Discovery in eHealth

2014-34 Christina Manteli (VU)
The Effect of Governance in Global Software Development: Analyzing Transactive Memory Systems

2014-35 Joost van Ooijen (UU)
Cognitive Agents in Virtual Worlds: A Middleware Design Approach

2014-36 Joos Buijs (TUE)
Flexible Evolutionary Algorithms for Mining Structured Process Models

2014-37 Maral Dadvar (UT)
Experts and Machines United Against Cyberbullying

2014-38 Danny Plass-Oude Bos (UT)
Making brain-computer interfaces better: improving usability through post-processing

2014-39 Jasmina Maric (UvT)
Web Communities, Immigration, and Social Capital

2014-40 Walter Omona (RUN)
A Framework for Knowledge Management Using ICT in Higher Education

2014-41 Frederic Hogenboom (EUR)
Automated Detection of Financial Events in News Text

2014-42 Carsten Eijckhof (CWI/TUD)
Contextual Multidimensional Relevance Models
2014-43 Kevin Vlaanderen (UU)
Supporting Process Improvement using Method Increments

2014-44 Paulien Meesters (UvT)
Intelligent Blauw. Met als ondertitel: Intelligence-gestuurde politiezorg in gebiedsgebonden eenheden

2014-45 Birgit Schmitz (OUN)
Mobile Games for Learning: A Pattern-Based Approach

2014-46 Ke Tao (TUD)
Social Web Data Analytics: Relevance, Redundancy, Diversity

2014-47 Shangsong Liang (UVA)
Fusion and Diversification in Information Retrieval

2015

2015-01 Niels Netten (UvA)
Machine Learning for Relevance of Information in Crisis Response

2015-02 Faiza Bukhsh (UvT)
Smart auditing: Innovative Compliance Checking in Customs Controls

2015-03 Twan van Laarhoven (RUN)
Machine learning for network data

2015-04 Howard Spoelstra (OUN)
Collaborations in Open Learning Environments

2015-05 Christoph Bösch (UT)
Cryptographically Enforced Search Pattern Hiding

2015-06 Farideh Heidari (TUD)
Business Process Quality Computation - Computing Non-Functional Requirements to Improve Business Processes

2015-07 Maria-Hendrike Peetz(UvA)
Time-Aware Online Reputation Analysis

2015-08 Jie Jiang (TUD)
Organizational Compliance: An agent-based model for designing and evaluating organizational interactions

2015-09 Randy Klaassen(UT)
HCI Perspectives on Behavior Change Support Systems

2015-10 Henry Hermans (OUN)
OpenU: design of an integrated system to support lifelong learning

2015-11 Yongming Luo(TUE)
Designing algorithms for big graph datasets: A study of computing bisimulation and joins

2015-12 Julie M. Birkholz (VU)
Modi Operandi of Social Network Dynamics: The Effect of Context on Scientific Collaboration Networks

2015-13 Giuseppe Procaccianti(VU)
Energy-Efficient Software
2015-14 Bart van Straalen (UT)
A cognitive approach to modeling bad news conversations
2015-15 Klaas Andries de Graaf (VU)
Ontology-based Software Architecture Documentation
2015-16 Changyun Wei (UT)
Cognitive Coordination for Cooperative Multi-Robot Teamwork
2015-17 André van Cleeff (UT)
Physical and Digital Security Mechanisms: Properties, Combinations and Trade-offs
2015-18 Holger Pirk (CWI)
Waste Not, Want Not! - Managing Relational Data in Asymmetric Memories
2015-19 Bernardo Tabuenca (OUN)
Ubiquitous Technology for Lifelong Learners
2015-20 Loïs Vanhé (UU)
Using Culture and Values to Support Flexible Coordination
2015-21 Sibren Fetter (OUN)
Using Peer-Support to Expand and Stabilize Online Learning
2015-22 Zhemin Zhu(UT)
Co-occurrence Rate Networks
2015-23 Luit Gazendam (VU)
Cataloguer Support in Cultural Heritage
2015-24 Richard Berendsen (UVA)
Finding People, Papers, and Posts: Vertical Search Algorithms and Evaluation
2015-25 Steven Woudenberg (UU)
Bayesian Tools for Early Disease Detection
2015-26 Alexander Hogenboom (EUR)
Sentiment Analysis of Text Guided by Semantics and Structure
2015-27 Sándor Héman (CWI)
Updating compressed column stores
2015-28 Janet Bagorogoza(TiU)
KNOWLEDGE MANAGEMENT AND HIGH PERFORMANCE; The Uganda Financial Institutions Model for HPO
2015-29 Hendrik Baier (UM)
Monte-Carlo Tree Search Enhancements for One-Player and Two-Player Domains
2015-30 Kiavash Bahreini(OU)
Real-time Multimodal Emotion Recognition in E-Learning
2015-31 Yakup Koç (TUD)
On the robustness of Power Grids
2015-32 Jerome Gard(UL)
Corporate Venture Management in SMEs
2015-33 Frederik Schadd (TUD)
Ontology Mapping with Auxiliary Resources
2015-34  Victor de Graaf (UT)
  Gesocial Recommender Systems

2015-35  Jungxao Xu (TUD)
  Affective Body Language of Humanoid Robots: Perception and Effects in Human Robot Interaction

2016

2016-01  Syed Saiden Abbas (RUN)
  Recognition of Shapes by Humans and Machines

2016-02  Michiel Christiaan Meulendijk (UU)
  Optimizing medication reviews through decision support: prescribing a better pill to swallow

2016-03  Maya Sappelli (RUN)
  Knowledge Work in Context: User Centered Knowledge Worker Support

2016-04  Laurens Rietveld (VU)
  Publishing and Consuming Linked Data

2016-05  Evgeny Sherkhonov (UVA)
  Expanded Acyclic Queries: Containment and an Application in Explaining Missing Answers

2016-06  Michel Wilson (TUD)
  Robust scheduling in an uncertain environment

2016-07  Jeroen de Man (VU)
  Measuring and modeling negative emotions for virtual training

2016-08  Matje van de Camp (TiU)
  A Link to the Past: Constructing Historical Social Networks from Unstructured Data

2016-09  Archana Nottamkandath (VU)
  Trusting Crowdsourced Information on Cultural Artefacts

2016-10  George Karafotias (VUA)
  Parameter Control for Evolutionary Algorithms

2016-11  Anne Schuth (UVA)
  Search Engines that Learn from Their Users

2016-12  Max Knobbout (UU)
  Logics for Modelling and Verifying Normative Multi-Agent Systems

2016-13  Nana Baah Gyan (VU)
  The Web, Speech Technologies and Rural Development in West Africa - An ICT4D Approach

2016-14  Ravi Khadka (UU)
  Revisiting Legacy Software System Modernization

2016-15  Steffen Michels (RUN)
  Hybrid Probabilistic Logics - Theoretical Aspects, Algorithms and Experiments

2016-16  Guangliang Li (UVA)
 社ocially Intelligent Autonomous Agents that Learn from Human Reward
2016-17  Berend Weel (VU)
Towards Embodied Evolution of Robot Organisms

2016-18  Albert Meroño Peñuela
Refining Statistical Data on the Web

2016-19  Julia Efremova (Tu/e)
Mining Social Structures from Genealogical Data

2016-20  Daan Odijk (UVA)
Context & Semantics in News & Web Search

2016-21  Alejandro Moreno Célleri (UT)
From Traditional to Interactive Playspaces: Automatic Analysis of Player Behavior in the Interactive Tag Playground

2016-22  Grace Lewis (VU)
Software Architecture Strategies for Cyber-Foraging Systems

2016-23  Fei Cai (UVA)
Query Auto Completion in Information Retrieval

2016-24  Brend Wanders (UT)
Repurposing and Probabilistic Integration of Data: An Iterative and data model independent approach

2016-25  Julia Kiseleva (TU/e)
Using Contextual Information to Understand Searching and Browsing Behavior

2016-26  Dilhan Thilakarathne (VU)
In or Out of Control: Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices, with Applications in Aviation and Energy Management Domains