Computer Recreations

ANDREW S. TANENBAUM
Wiskundig Seminarium, Der Vrije Universiteit, Amsterdam-11

A Heuristic for Playing Jotto*

Jotto is a popular word game for two players. It is of interest here because it unquestionably requires some intellectual ability for people to play it well, and it is a game in which a simple program can beat most human players nearly all the time.

RULES OF JOTTO

Before the game starts, each player chooses a five-letter English word as his secret word. The players alternate asking each other five-letter test words, one word per turn. The first player to guess his opponent’s secret word wins.

If your opponent’s test word is your secret word, you just admit it and the game is over. Otherwise you tell him how many of his test word’s letters are contained in your secret word. If his test word and your secret word have no letters in common, the answer is 0. If both words contain the same five letters, the answer is 5. If your secret word contains a repeated letter and his test word contains that letter only once (or vice versa) it counts as one. If both the test word and the secret word contain two occurrences of the same letter, it counts twice. The number of letters the two words have in common is called the number of jots. A few examples are shown in Table I.

<table>
<thead>
<tr>
<th>Secret word</th>
<th>Test word</th>
<th>Number of jots</th>
<th>Matching letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>HORSE</td>
<td>CLUMP</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>HORSE</td>
<td>NAILS</td>
<td>1</td>
<td>S</td>
</tr>
<tr>
<td>HORSE</td>
<td>SOAPY</td>
<td>2</td>
<td>OS</td>
</tr>
<tr>
<td>HORSE</td>
<td>STORK</td>
<td>3</td>
<td>ORS</td>
</tr>
<tr>
<td>HORSE</td>
<td>RESET</td>
<td>3</td>
<td>RSE</td>
</tr>
<tr>
<td>HORSE</td>
<td>EASES</td>
<td>2</td>
<td>SE</td>
</tr>
<tr>
<td>SLEEP</td>
<td>PEALS</td>
<td>4</td>
<td>SLEEP</td>
</tr>
<tr>
<td>SLEEP</td>
<td>PEELS</td>
<td>5</td>
<td>SLEEP</td>
</tr>
<tr>
<td>ERROR</td>
<td>ROTOR</td>
<td>3</td>
<td>RRO</td>
</tr>
<tr>
<td>ERROR</td>
<td>MOTOR</td>
<td>2</td>
<td>RO</td>
</tr>
</tbody>
</table>

Note that the position of the matched letters within the secret word is unimportant, i.e. no distinction is made between a letter that matches the corresponding letter in the secret word, and a letter that matches some other letter in the secret word. Also note that to win, a

* N° is delighted to print this article.

Received 16 July 1973

© 1973 by John Wiley & Sons Ltd.,

397
player must guess his opponent's secret word, not just get a score of 5 jots. For example, if
the first player asks LEAST as a secret word and is told that it contains 5 jots, he may not
necessarily have won, since his opponent's secret word might be STALE or STEAL or
TALES or SLATE. Furthermore, it is important to realize that only genuine five-letter
English words may be used as test words and secret words. Jargon, abbreviations, acronyms
and foreign words are forbidden. If an arbitrary combination of five letters were allowed,
the game would degenerate into a mere combinatorial problem.

THE STRATEGY USED BY PEOPLE

Because the heuristic to be described is completely different from the way human beings
play jotto, the human approach will be sketched first. People generally attempt to deduce at
least 3 or 4 letters one at a time before trying to guess the secret word. For example,
player 1 might start off with STONE as his test word and be told that it contains 1 jot. On
his next turn he might ask CLUMP and be told that it has 0 jots. He can now use the fact
that the letters C, L, U, M and P do not occur in his opponent's secret word to determine
which of the letters in STONE is the one in the secret word. On his third turn he might ask
STUMP. If the reply is 1, the letter is either S or T. If the reply is 0, the letter is either
O, N or E. Assuming a reply of 1, he might try LUMPS next. A reply of 1 means the secret
word contains an S and a reply of 0 means it contains a T. Having discovered which letter
STONE contains, the player probably will, on his next turn, ask a test word all of whose
letters are distinct from those in STONE and CLUMP (e.g. HAIRY) to begin the
elimination process again.

Human players tend to pick secret words containing multiple repetitions of a letter or
words in which several permutations of the letters are legal words. A few of my favourite
words are (were) SEXES, FUZZY, ERROR, QUAFF and TEALS.

THE STRATEGY USED BY THE PROGRAM

The jotto program described below (which runs on a CDC Cyber 73) uses an approach which
requires repeatedly searching a large dictionary, a task at which computers are noticeably
better than human beings.

Before each game starts, the program reads in an extensive dictionary of five-letter
English words. The dictionary contains about 6,500 words, ranging from ABACK to
ZOOMS. The program chooses its secret word from this list at random. On its turn the
program chooses a test word, and types the word at its human opponent's on-line terminal,
inquiring how many jots it contains. After being told how many jots there are in the test
word, for example N, the program proceeds as follows. It examines the first word in the
dictionary, initially ABACK, to determine the answer to the question: 'If ABACK were
his secret word, would he have said that my last test word contained N jots?' This question
can be trivially answered by comparing ABACK to the test word. If the answer is no,
ABACK could not possibly be his secret word, so it is removed from the dictionary.

Each word still in the dictionary is examined in turn and removed if, considered as the
secret word, it would have caused a response other than N to the test word. If the test word
received a score of 5 jots, the program asks if that word is the secret one. If not, the test
word is also removed from the dictionary. As the game proceeds, the dictionary becomes
smaller and smaller. Eventually one of three things will happen: the human player will
guess the machine's secret word and win; the machine will guess the human player's secret
word and win, or the program will eliminate the entire dictionary before either side has won. The latter occurs if the person's secret word is not in the program's dictionary (an exceedingly rare occurrence) in which case the machine gracefully concedes defeat, or if the person makes a mistake in replying to some test word (a fairly common occurrence).

Even if the program chooses its test words from the list of remaining words at random, it performs extremely well, averaging only 8.6 turns to guess its opponent's secret word. Few people can play this well consistently. Furthermore, the program can be improved by having it choose its test words more wisely. One useful technique is minimax.

Before choosing a test word, the program (randomly) selects a list of candidate test words to be evaluated, the best of which will be used as the test word. Each candidate test word is evaluated as follows. For each of the six possible responses to it, 0, 1, 2, 3, 4 or 5, the number of words eliminated from the dictionary by the test word is computed. The smallest of these six numbers is the 'worst case' score, the number of words the program can be guaranteed to eliminate if it chooses this candidate as the test word. The worst case score is computed for each candidate test word, and the one with the highest score (i.e. the best guarantee) is chosen.

Empirical testing indicates that the program's performance improves as the list of candidate test words is lengthened, up to 25 words. Above this, further increases have no effect. With 25 candidate test words the program averages 7.4 turns to guess its opponent's secret word.

It is interesting to note that words containing several unusual letters, although difficult for human beings are especially easy for the program. Empirical studies show, for example, that the jotto program can guess FUZZY in an average of 6.1 turns. In 23 per cent of the games the program guessed FUZZY on or before the fifth turn, a remarkable feat. When the program gets 0 as a reply to a test word, it eliminates all words containing any of the letters in the test word. If these are common letters, many words will be eliminated. Secret words such as FUZZY increase the probability that the test word will get 0 as a reply and as a consequence, the dictionary dwindles more rapidly than usual.

ACKNOWLEDGEMENT

The author would like to thank Roy Carlson for helping with the programming.
Erratum

CDC 6000-SERIES REGISTER SAVE/RESTORE
(Article in Software—Practice and Experience, 2, 377–387, 1972)
by
JOHN L. BAKER

The ‘restore’ column of Table I should read:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Space</td>
<td>40</td>
</tr>
<tr>
<td>Minimum time</td>
<td>125.9</td>
</tr>
<tr>
<td>Expected time</td>
<td>126.0</td>
</tr>
<tr>
<td>Maximum time</td>
<td>137.1</td>
</tr>
</tbody>
</table>

Page 382, lines five and six should be interchanged so as to read:

SA1   B3REG
SA2   B4REG

Page 382, line 14 should read:

* RESTORE A1:A7, X0, X2:X5

This line should be followed immediately by two new additional lines reading:

SA5   X0REG
BX0   X5

Page 385, line five from the bottom should read:

SX6   X6+B7   EXTEND X6[17] AS SIGN