NOTES

COMPRESSING A CHESS-ENDGAME DATABASE

Christopher A. Marris
Hull, England

ABSTRACT

A method of compressing a chess-endgame database taking advantage of the arrangement of the number of moves-to-mate between adjacent positions within the database is described. A subset of KQK with White to move (WTM) is used to illustrate the technique.

1. INTRODUCTION

One of the problems with computer-chess endgame databases (especially when microcomputer-implemented) is their large size. To show how such databases may profitably be compressed considerably, we take as an example a subset of the KQK endgame. The subset considered has fixed positions for the two Kings (WK on e8; BK on a8). Diagram 1 shows the number of moves to mate for each of the 45 legal positions.

`Diagram 1`

`x` = forbidden position of the white Queen.

We number the rows and files from 0 to 7, and use the following abbreviations: WKR = row(WK); WKF = file(WK); BKR = row(BK); BKF = file(BK); WQR = row(WQ); WQF = file(WQ). A possible indexing function F1 for each KQK position may now be:

\[ F1 = 32768 \times \text{BKR} + 4096 \times \text{BKF} + 512 \times \text{WKR} + 64 \times \text{WKF} + 8 \times \text{WQR} + \text{WQF}. \]

This indexing method guarantees that each position in the KQK database has a different entry point in the database, given by F1.
2. DIFFERENT INDEXING FUNCTIONS

The subset of moves-to-mate (M) of the KQK database, corresponding with the white King and black King confined to positions e8 and a8, respectively, using FI as indexing function, is given in Figure 1 (values of 0 indicate illegitimate KQK positions, otherwise moves to mate).

\begin{verbatim}
entry[233216]....................................................................................................................................entry[233279]
\end{verbatim}

Figure 1: Moves-to-mate (M) using FI as the indexing function.

If we define an M-interval as a contiguous set of database entries with the same M-values (disregarding illegitimate positions), the information in Figure 1 may equally well be expressed by Table 1.

<table>
<thead>
<tr>
<th>M-interval</th>
<th>M-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>entry[233216]</td>
<td>3</td>
</tr>
<tr>
<td>entry[233217]</td>
<td>4</td>
</tr>
<tr>
<td>entry[233218]</td>
<td>3</td>
</tr>
<tr>
<td>entry[233219]</td>
<td>4</td>
</tr>
<tr>
<td>entry[233220]</td>
<td>3</td>
</tr>
<tr>
<td>entry[233221]</td>
<td>4</td>
</tr>
<tr>
<td>entry[233222]</td>
<td>3</td>
</tr>
<tr>
<td>entry[233223]</td>
<td>4</td>
</tr>
<tr>
<td>entry[233224]</td>
<td>3</td>
</tr>
<tr>
<td>entry[233225]</td>
<td>4</td>
</tr>
<tr>
<td>entry[233226]</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 1: M-intervals corresponding to Figure 1.

Note that by convention we include illegitimate positions as part of the next-higher indexed legitimate position.

That the indexing function may have a profound influence on the number of intervals generated is shown by exhibiting the alternative indexing function F2:

\[ F2 = 32768 \times BKR + 4096 \times BKF + 512 \times WKR + 64 \times WKF + 8 \times WQF + WQR. \]

The above procedure is repeated to obtain Figure 2 and Table 2.

\begin{verbatim}
entry[233216]....................................................................................................................................entry[233279]
\end{verbatim}

Figure 2: Moves-to-mate (M) using F2 as the indexing function.
Table 2: M-intervals corresponding to Figure 2.

So, whereas indexing function F1 leads to 10 M-intervals, F2 reduces this figure (by effectively rearranging the KQK subset database) to an optimal value of 2. (Optimal because the number of M-intervals equals the number of different M-values.)

Other indexing functions may be used. The number of different indexing functions of the kind of F1 and F2 equals 720 (being 6!, i.e., all permutations of the 6 parameters (BKR, BKF, etc.) involved), but of course also other types of indexing functions may be tried.

For the complete KQK database with WTM the maximum number of moves-to-mate is known to be 10. It is tempting to devise an indexing function yielding, for the complete KQK database, a number of M-intervals which exceeds the required minimum of 10 by the lowest amount possible.

3. REFERENCES


