RETROGRADE ANALYSIS OF CERTAIN ENDGAMES

Ken Thompson

AT&T Bell Laboratories Murray Hill, New Jersey, U.S.A.

Introduction

Computers have been able to exhaustively solve certain simple chess endgames. The general method is to work backwards from mates (or known winning positions). Wins-in-one are marked by generating predecessor positions to mates. Wins-in-two are then marked by generating predecessor positions to mates and wins in one, etc. When no more winning positions are generated, the unmarked positions are either illegal, draws or losses.

The major complexity of this problem is the number of pieces on the board. A minor complexity is the existence of a pawn, which complicates the symmetry. The first work in this field¹ consisted of analysis of one and two-piece (not counting Kings) endgames. Since that time there has been much material published on endgames with less than three pieces. This paper describes work on three-piece endgames. There has been some prior work on three-piece endgames.^{2,3} There is also a good recent bibliography to the field.⁴

The Gödel Function

The basis of a retrograde analyzer is a subprogram that converts a chess position into a unique Gödel number (G) that is used to index a database of positions. The inverse transformation (from G to position) is also important. Different transformations are used for positions with and without pawns. The existence of a pawn destroys one of the symmetries enjoyed with pure pieces. Let me first describe the pure-piece transformations.

The canonical position has two Kings and three pure pieces. The White King is confined to the "octant" defined by the squares a8-d8-d5-a8. This is not a restriction since any position can be rotated and reflected to place the White King in this area. The Black King encoding is based on the position of the White King. In particular, the Black King cannot be next to the White King and if the White King is on the a8-d5 diagonal, the Black King can be rotated to be above that diagonal. The King encodings are implemented with a table lookup. This lookup also selects what rotation/reflection to apply to the other three pieces. There are 462 legal positions of two Kings where the described symmetries are removed. The other three pieces are encoded by their position on the board after the rotation/reflections indicated by the King positions. If the piece is missing (captured) then it is assigned the position of the White King, which would otherwise be illegal. Thus the encoded G has the following parts:

0-461 Position of two Kings

0-63 Position of Piece 1.

0-63 Position of Piece 2.

0-63 Position of Piece 3.

giving G a range of $462 \times 64 \times 64 \times 64$ or 121,110,528.

There are potentially other symmetries that could be exploited depending on what pieces are on the board. For example, if there are two White Rooks, they could exchange places without altering the position. More notably, if there are two White Bishops, and it is assumed that they are of opposite color, each could be assigned a 32 square subset of the board rather than 64 squares. None of these potential symmetries were exploited.

The Method

The retrograde analysis is accomplished by 4 programs: P_1 thru P_4 . These programs operate on files of sets of chess positions. Each file is a bit map of 121,110,528 bits. If a bit is on in a file then the corresponding G encoded chess position is considered in the set of chess positions represented by the file. There are 5 files that are successively manipulated by the programs. File W is a list of all currently known White-to-move and win positions. B is all currently known Black-to-move and lose positions. W_i is the latest newly found White-to-move and win in i moves. This file is added (logically ored) to W to bring W up to date. B_i is the latest Black-to-move and lose in i moves. The remaining file, J_i is a temporary file that is a superset of B_i as described below.

Program P_1 is for initialization. It is run exactly once at the beginning and creates the file B_0 , those positions where Black-to-move and Black is mated. P_1 loops through all G positions, converts each to chess-board representation and examines each for a legal mate. If it is mate, then the corresponding bit is set in file B_0 . B is initialized to B_0 . W is initialized to all zeros and i is set to zero.

Program P_2 examines each position in file B_i and for each position, generates all possible legal predecessor positions. These positions have White-to-move and with at least one White move can obtain a B_i position. They are therefore the new W_{i+1} positions if they have not been found before — that is if they are not in W. After P_2 finishes reading B_i , consulting W and creating W_{i+1} , W_{i+1} is added to W.

Predecessor positions are formed by an un-move generator. This is the same as a move generator but a) it is illegal to start in check, but legal to un-move into check and b) it is illegal to capture, but legal to un-capture by leaving an enemy piece behind. In the games that we are discussing, un-castle and un-enpassant are not encountered. Un-promotion is described later under pawn endgames.

Program P_3 is exactly the same as P_2 except with Black-to-move. P_3 reads W_{i+1} , generates Black predecessor positions, consults B, and creates J_{i+1} . The positions in J_{i+1} are Black-to-move and lose if Black wanted to mate himself. Of these positions, only those that can be forced to a winning White position are losses.

This brings us to P_4 that reads J_{i+1} , generates Black successor positions and examines each in W. If all such successors are in W, then that position is added to B_{i+1} . When P_4 finishes, the new B_{i+1} positions are added to B. The process is iterated by incrementing i and repeating programs P_2 , P_3 , and P_4 in turn until no more positions are generated.

When the process finishes, W contains all positions in which White can force a win, with White-to-move. These positions are partitioned into the files W_i ; White-to-win in i, with best play. The easiest way to save the results in a database is to make a file with 121,110,528 bytes. Byte j contains i if bit j is set in file W_i . In other words, byte j in the result file is the number of moves for White to win, and zero if the position is not a win for White. This result file is incrementally updated after the generation of each W_i file. The total space required is then the result file, W, B, and only two active files from B_i , W_i and J_i . This totals $12 \times 121,110,528$ bits of secondary storage, or about 175 megabytes.

Results

These programs were implemented on a Sequent Balance 8000 computer. This computer consists of 12 National 32032 microprocessors on a 16 megabyte shared memory. The programs were designed to divide the work by assigning each processor every Nth position. The disk traffic was drastically reduced by allocating a megabyte of real memory per processor as a cache on the files that were randomly accessed. The programs were run in the background with an average of four processors working simultaneously. A typical pure-piece endgame would be solved in two to three weeks of real time.

Some of the results tabulated below are labeled "Max to Mate." These games were solved in one pass as described above. Max refers to the number of moves to mate with best play.

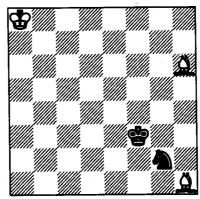
The other positions were solved in two passes. These are labeled "Max to Win." In these endgames, the first pass solved all subgames with fewer than the full complement of pieces to mate. The second pass then used these positions as B_0 along with all mates to obtain results with the complete complement of pieces. Thus the two-pass, "Max to Win," method solves positions with an objective function of either mate or capture into a won sub-game. This is precisely the conditions for the 50-move rule.

The "Percent Wins" column of Table 1 represents the number of positions that are wins divided by the total number of legal positions with White to move. These numbers will vary depending on the Gödel function since some positions are counted more than once. Also note that it is White to move in what is essentially a random position. This is a large advantage and the win percentage favors White. It is hard to characterize, but a win percentage of about 40 is indicative of a drawn endgame and a percentage of about 90 is a won endgame.

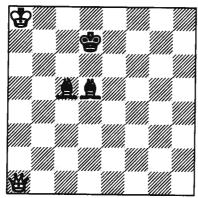
		1,7	3.4	ъ .
White	Black	Max	Max	Percent
		to Mate	to Win	Wins
۵۵	Ð		66	91.8
*	22		63	89.7
&	22	ļ	42	93.1
유 유 주	စ္တစ္အ		71	92.1
国	耳		33	35.9
里夏	耳		59	40.1
異国	買	31	1	94.3
耳骨	旦	35		95.9
4 0	&	41		48.4
₩2	*	33]	53.4
4日	₩	67		92.1
쌉쌉	₩	30	İ	94.0

Table 1

The first group of endgames in Table 1 were solved for chess interest. All four of the endgames in this group are considered to be drawn. This work shows that they are really wins. The next two groups were done as terminal promotion positions for pawn endgames. Some extreme cases are illustrated by the following "best play" examples.

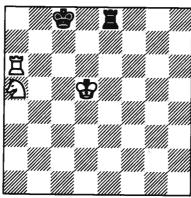


Initial Position



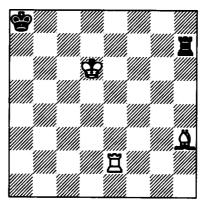
Initial Position

1 \$\text{\$\psi\$} \text{\$\psi\$} \text{\$\psi



Initial Position

1 旦a8+ 含d7 2 旦a7+ 含c8 3 含d6 旦d8+ 4 含c6 含b8 5 旦b7+ 含a8 6 旦h7 旦c8+ 7 含b6 旦b8+ 8 含c5 旦g8 9 旦h4 旦b8 10 公c6 旦b2 11 旦h7 旦c2+ 12 含d6 旦d2+ 13 含c7 旦h2 14 旦d7 旦d2 15 公d4 旦b2 16 含c6 旦b7 17 旦d5 旦b4 18 公b5 旦c4+ 19 含b6 含b8 20 旦e5 旦c1 21 旦e8+ 旦c8 22 旦e1 旦c2 23 公d4 旦b2+ 24 含c6 含a8 25 旦f1 旦b4 26 公b5 旦c4+ 27 含b6 含b8 28 公d6 旦b4+ 29 含c6 含a8 30 旦f8+ 旦b8 31 公c8 旦b3 32 公b6+ 含a7 33 旦a8#



Initial Position

1 点f5 闰h4 2 点d3 闰f4 3 点e4+ 含a7 4 点c6 闰g4 5 含c7 闰g7+ 6 点d7 闰g6 7 点e6 闰g7+ 8 含c6 闰g1 9 闰a2+ 含b8 10 闰b2+ 含a8 11 含b6 闰c1 12 点f5 闰c3 13 闰b1 含b8 14 闰b4 闰a3 15 点d7 闰a2 16 闰h4 闰b2+ 17 点b5 闰c2 18 点c4 闰b2+ 19 含c6 闰f2 20 闰h8+ 含a7 21 闰h7+ 含b8 22 闰b7+ 含a8 23 闰b4 闰g2 24 点d3 闰g3 25 闰d4 ቯf3 26 点c4 闰h3 27 闰d8+ 含a7 28 点d5 闰h2 29 闰d7+ 含b8 30 闰b7+ 含a8 31 闰b1 闰c2+ 32 含b6+ 含b8 33 点e6 闰d2 34 含c6+ 含a7 35 闰a1+ 含b8 36 点d5 闰h2 37 闰b1+ 含a7 38 点e4 闰h6+ 39 含c5 闰b6 40 闰h1 闰a6 41 闰h8 闰a5+ 42 含c6 闰g5 43 闰h7+ 含a6 44 点d5 含a5 45 含c5 闰g6 46 闰h2 闰g4 47 闰b2 闰h4 48 闰b7 闰h6 49 点f7 闰f6 50 点c4 闰f5+ 51 点d5 闰f6 52 闰b5+ 含a6 53 闰b2 含a7 54 闰b7+ 含a6 55 闰e7 含a5 56 点e6 含a6 57 点c8+ 含a5 58 딜a7+ 딜a6 59 딜×a6#

And with a Pawn

When one of the pieces is a pawn, everything gets harder. The Gödel function loses symmetries; the pawns can promote into sub-games that must be solved independently; each pawn position must be treated as a sub-game and solved independently; and there are numerous smaller annoyances. The overall structure remains. First all promoted sub-games are solved. These are all combined to create the file W_{8^*} , White-to-move, pawn on the 8th, and win in any number of moves. Second, a new program P_5 makes pawn un-moves to create the file B_{70} . These positions are combined with mates on the 7th and also wins without the pawn. Programs P_2 , P_3 and P_4 are then iterated to create files W_{7^*} , B_{7^*} and J_{7^*} . Again, this is done in two passes to first create wins without the full complement and second to generate all wins. And then P_5 is run to back up W_{7^*} into B_{60} . All goes well until rank 2 when, because of the initial pawn move, files W_{3^*} and W_{4^*} must be combined before creating B_{20} .

The basic size of the G function is 83,886,080 to accommodate the pawn on one of the four files and one rank at a time. The total storage required is about 120 megabytes per rank. After the six result files of 83 megabytes each are complete, they are inverted into four rank-oriented files of 117 megabytes each. It takes about six weeks real time to create a complete pawn data base, not counting the time to create the promotion sub-games.

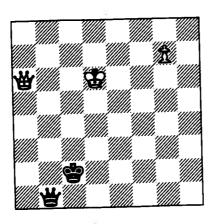
쌓î versus 쌀

The first pawn endgame attempted was Queen and Pawn vs. Queen. Table 2 below summarizes the results "max to win" and "percent wins" for each of the 24 initial pawn positions. Note that "max to win" means moves until capture, mate or pawn move; precisely the 50-move definition. Unfortunately, a clerical error surfaced. The under-promotions were accidentally discarded and so the results are for Queen promotion only. The chess literature contains some Rook under-promotions but every one that was examined could be won in a few more moves with Queen promotion. So the max figures for the 7th rank may have small errors, but the other figures in table 2 are probably correct.

				12 25 (
l 7	70 84.0	55 8 4 .7	43 85.4	42 85.6
6	71 71.0	61 76.5	46 79.3	58 <i>7</i> 7.6
5	33 57.2	38 63.8	43 72.7	45 70.1
4	29 51.8	30 55.6	48 67.2	64 65.6
3	20 48.5	51 52.7	53 61.5	54 58.3
1 -	17 48.6	31 52.7	47 62.7	41 59.0
2	1/ 40.0	31 32.7	- X/ OZ./	1
1	a	b	C	<u>a</u>

Table 2

There is one QPvQ example in Komissarchik and Futer. That analysis is presented below annotated with the current data base. The numbers in parentheses are the discrepancies in number of moves to mate, capture or promotion.



Initial Position

1 ... 쌓b4t 2 \$\frac{1}{2}\$ e6 \(\frac{1}{2}\$ g4t 3 \) \$\frac{1}{2}\$ f6 \(\frac{1}{2}\$ f4t 4 \) \$\frac{1}{2}\$ g6 \(\frac{1}{2}\$ e4t 5 \) \$\frac{1}{2}\$ g5 \(\frac{1}{2}\$ e4t 6 \) \$\frac{1}{2}\$ f5 \(\frac{1}{2}\$ f5 \) \$\frac{1}{2}\$ f6 \(\frac{1}{2}\$ d4t 10 \) \$\frac{1}{2}\$ f7 \(\frac{1}{2}\$ d7t 11 \) \$\frac{1}{2}\$ g6 \(\frac{1}{2}\$ g4t 12 \) \$\frac{1}{2}\$ h7 \(\frac{1}{2}\$ h3t 13 \) \$\frac{1}{2}\$ g8 \(\frac{1}{2}\$ f5 14 \) \$\frac{1}{2}\$ a2t \(\frac{1}{2}\$ c1 15 \) \$\frac{1}{2}\$ h2 \(\frac{1}{2}\$ d5t 16 \) \$\frac{1}{2}\$ h8 \(\frac{1}{2}\$ d4 17 \) \$\frac{1}{2}\$ c7 \(\frac{1}{2}\$ b1 18 \) \$\frac{1}{2}\$ h7 \(\frac{1}{2}\$ d5t 19 \) \$\frac{1}{2}\$ b6 \(\frac{1}{2}\$ d5t 23 \) \$\frac{1}{2}\$ d7 \(\frac{1}{2}\$ d5t 26 \) \$\frac{1}{2}\$ c8 \(\frac{1}{2}\$ d5t 27 \) \$\frac{1}{2}\$ b8 \(\frac{1}{2}\$ d5t 28 \) \$\frac{1}{2}\$ a2 \(\frac{1}{2}\$ d5t 29 \) \$\frac{1}{2}\$ b6 \(\frac{1}{2}\$ d5t 30 \) \$\frac{1}{2}\$ a6 \(\frac{1}{2}\$ a2t 31 \) \$\frac{1}{2}\$ a5 \(\frac{1}{2}\$ g8 32 \) \$\frac{1}{2}\$ b4t \(\frac{1}{2}\$ a2 33 \) \$\frac{1}{2}\$ d4 \) \$\frac{1}{2}\$ e6t

33 ... **當c8**t (1)

34 含b5 쌓e8t 35 含b4 쌓b8t

35 ... **皆**e1† (5)

36 එc3 එg3t 37 එd2

37 \$c2 (10) 37 \$c4 (8) 37 \$b4 (3)

37 ... 谐g2t 38 含e1 谐h1t 39 含f2 谐h2t 40 含f3 谐h3t 41 含f4 谐h2t 42 含g5 谐g3t 43 含f6 谐f3t 44 含e6 谐c6t 45 含e5 谐e8t 46 含f4 谐f7t 47 含g3 谐g6t 48 含h3 谐h7t 49 含g2 谐g6t 50 含f1 谐b1t 51 含e2 谐b5t 52 含d2 谐b3 53 谐a7t

53 皆d3 (1)

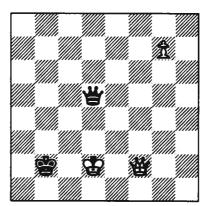
53 ... **\$**b2 54 **\$**f2 **\$**g8

54 ... *d5† (1)

55 🗳 b 6 † 🕏 a 3 5 6 🗳 b 7 🕏 a 4 5 7 🕏 c 3 🕏 a 5 5 8 🗳 b 4 † 🕏 a 6 5 9 🗳 c 4 †

The one-move discrepancies near the end can be explained by under-promotions. For example:

а



Position after 54 ... "d5+

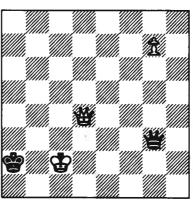
The data base play is:

55 含e3+ 含c3 56 발f3 발g8 57 含f2+ 含d4 58 발f8 발a2+ 59 含g3 발b3+ 60 含h4 etc.

But this can be improved by:

55 含e3t 含c3 56 含f6t 含c4 57 含f4t 含c3 58 含c7t 含b2 59 含b8t 含a1 60 g8티!

The discrepancies at moves 33 and 35 are simply to avoid the position at move 37. Thus the only real dispute is the position at move 37:



Position after 37 &c2

Komissarchik and Futer imply that this position is a win in at least 23 moves. The current work claims that it is a win in 13. The following analysis is given in support of the current work.

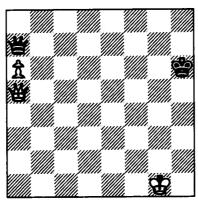
- 37 ... "g2+? 38 "gd2 "gc6+ 39 gd3+ and promote next.
- 37 ... 쌓c7t? 38 쌓c3 쌓h2t 39 쌓d2 쌓h7t 40 쌓c3t and mate.
- 37 ... 쌀g6+! 38 \$d1 쌓e6? 39 쌀a7+ \$b2 40 쌀b8+ \$a1 41 쌀a8+ \$b1 42 g8 쌓or more simply
 - 41 g8耳!
 - 38 ... 쌀h5t? 39 쌓c1 쌀h1t 40 쌓d1
 - 38 ... 쌓f7! 39 \$d2 \$b3 40 \$e1 *e6+ 41 \$f2 *bf7+ 42 \$g3 *c7+ 43 \$g2 *b7+ 44 \$g1 *bf7 45 *bg4 *bg8 46 *bf3+

39 ... \$\dagga a3 40 \$\dagga g8? 41 \$\dagga f3\tau a4 2 \$\dagga b7\$ and Black will have to leave the pawn to prevent mate.

40 ... 쌓f2+ 41 含d3 쌓f1+ 42 含e3 쌓e1+ 43 含f3 쌓d1+ 44 含g3 쌓d6+ 45 含g2 쌓d2+ 46 含g1 쌓c1+ 47 含h2 쌓c2+ 48 쌓g2 쌓h7+ 49 쌓h3+

September 1986

The author would welcome any input which helps resolve this discrepancy. The following is the maximal QPvQ play:



Initial Position

1 ውg2 ውg7 + 2 ውh1 ውf6 3 ውd2+ ውg7 4 ውd7ተ ውg6 5 ውd3ተ ውg7 6 ውg3ተ ውf8 7 ውb8ተ ውf7 8 ውc7ተ ውg6 9 ውc4 ውf3ተ 10 ውh2 ውe3 11 ውf1 ውe5ተ 12 ውh1 ውh8ተ 13 ውg2 ውa8ተ 14 ውg1 ውa7ተ 15 ውh1 ውd7 16 ውf2 ውh5 17 ውe2ተ ውg6 18 ውe4ተ ውh5 19 ውc4 ውa7 20 ውe2ተ ውg5 21 ውg2 ውd4 22 ውf2 ውe4ተ 23 ውf3 ውd4 24 ውe2 ውg6 25 ውb5 ውe3 26 ውf1 ውe4 27 ውf2 ውd4ተ 28 ውf3 ውh6 29 ውc6ተ ውg7 30 ውc7ተ ውg6 31 ውg3ተ ውh5 32 ውh3ተ ውg5 33 ውe6 ውd1ተ 34 ውe2 ውd5ተ 35 ውf2 ውf5ተ 36 ውg1 ውb1ተ 37 ውg2 ውg6 38 ውc4 ውh5ተ 39 ውf2 ውf6ተ 40 ውe3 ውe5ተ 41 ውd3 ውg5 42 ውc2 ውe3 43 ውb2 ውg6 44 ውb5 ውh6 45 ውc2 ውe6 46 ውc3 ው66 47 ውc4 ውg7 48 ውg5ተ ውh8 49 ውh5ተ ውg7 50 ውg4ተ ውf7 51 ውf5ተ ውg7 52 ውc8 ውf4ተ 53 ውb5 ውf6 54 ውc6 ውb8 ውb2ተ 55 ውc5 ውf2ተ 56 ውd6 ውg6 57 ውd5ተ ውh7 58 ውc7ተ ውh8 59 ውc3ተ ውh7 60 ውc6 ውf5 61 ውb6 ውe6ተ 62 ውb7 ውe4ተ 63 ውb8 ውb1ተ 64 ውc7 ውa2 65 ውc6 ውh8 66 ውb8 ውb3ተ 67 ውb7 ውg3ተ 68 ውa8 ውh3 69 ውc6 ውg4 70 ውc3ተ ውh7 71 a7

耳介 versus 耳

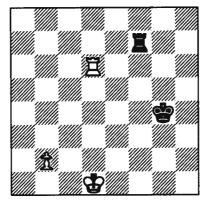
Table 3 is for Rook and Pawn vs Rook.

7	20 88.1	24 88.6	20 88.8	23 89.1
6	28 73.0	19 75.1	20 75.1	20 75.8
5	20 61.7	19 66.6	18 65.9	20 65.9
4	27 54.2	18 62.1	19 60.7	20 59.0
3	25 45.4	30 56.4	24 55.0	32 50.4
2	25 44.0	35 56.1_	33 53.8	33 48.7
	a	b	С	d

Table 3

There is one RPvR example in Arlazarov and Futer. It is not possible to compare that to the current work because of differences in objective function. The goal in Arlazarov and Futer is to promote, while in the current work, the goal is to push the pawn to the next square.

The following is the longest RPvR variation:

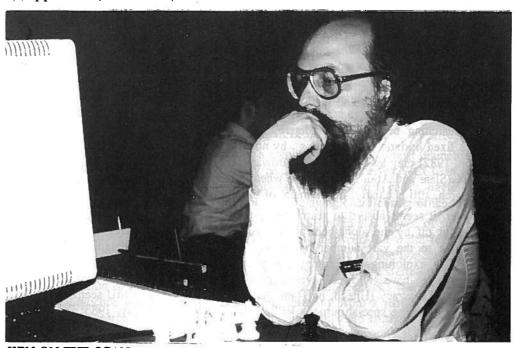


Initial Position

1 旦d5 旦f2 2 含c1 含f4 3 旦d2 旦f1+ 4 旦d1 旦f2 5 旦e1 旦g2 6 旦e8 旦h2 7 含b1 旦h7 8 含c2 旦c7+ 9 含d3 旦d7+ 10 含c3 旦c7+ 11 含d4 旦b7 12 旦f8+ 含g4 13 含c3 旦c7+ 14 含d3 旦b7 15 含c2 旦c7+ 16 含b1 旦h7 17 旦f6 旦b7 18 旦f2 旦b8 19 含c1 含g5 20 含c2 旦c8+ 21 含d2 旦b8 22 含c3 旦c8+ 23 含b4 旦b8+ 24 含a5 旦a8+ 25 含b6 旦b8+ 26 含a7 旦b5 27 含a6 旦b8 28 旦c2 含f6 29 旦c6+ 含e5 30 旦b6 旦a8+ 31 含b5 含d4 32 旦d6+ 含e5 33 旦c6 旦b8+ 34 旦b6 旦a8 35 b4

References

- 1. T. Ströhlein, "Untersuchungen über kombinatorische Spiele," Dissertation, Fakultät für Allgemeine Wissenschaften der Technischen Hochschule München (1970).
- 2. Z. A. Komissarchik and A. L. Futer, "Ob Analize Ferzevogo Endshpilia pri Pomoshchi EVM," *Problemy Kybernet*, Moscow **29**, pp. 211-220 (1974).
- 3. V. L. Arlazarov and A. L. Futer, "Computer Analysis of a Rook Endgame," pp. 361-371 in *Machine Intelligence 9*, ed. J. E. Hayes, D. Michie, and L. J. Mikulich, Ellis Horwood Limited, Chichester, England (1979).
- 4. H. J. van den Herik and I. S. Herschberg, "A Data Base on Data Bases," ICCA Journal 9(1), pp. 29-34 (March 1986).



KEN ON THE SCAN

Five digits for digitally contemplating five-piece endgames.

Photo by M.T. Fürstenberg