

A Taxonomy of Concepts for Evaluating Chess Strength

Hans Berliner,

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA. 15213-3890,

Danny Kopec, and Ed Northam

Department of Computer Science
University of Maine, Neville
Orono, Maine, 04469-0122

Abstract

This paper is an attempt to classify that which is important in chess. Other authors have chosen to make classifications relating to, for example, the material on the board, or whether there exists an attack on the king. We find that this is sufficient for an introductory treatment of the subject of advantages. However, the chess specialist understands that two apparently similar positions may, in fact, be quite different since there may be a very important difference in the features of each.

Earlier attempts to develop test suites for evaluating human and computer chess strength [12,8], although valuable, have had clear drawbacks in terms of their depth, range and number of positions examined. We present a taxonomy of positions in chess that require special knowledge. The taxonomy is what drives our selection of positions, and not vice-versa. It is easy to understand what a passed pawn is, and a bit of classic advice such as "Passed pawns must be pushed" makes use of the simple metric that a passed pawn becomes more valuable with each advance. However, there are outside passed pawns, protected passed pawns, blockaded passed pawns, and passed pawn masses. Each requires its own understanding, and frequently other features of a position can cause great variation in what may at first appear to be positions that should be treated very similarly.

This research was partly sponsored by the Defense Advanced Research Projects Agency (DOD), ARPA Order No. 4976, Amendment 20, monitored by the Air Force Avionics Laboratory Under Contract F33615-87-C-1499.

Our taxonomy is per force preliminary, as a work of this kind is essentially never complete. However, this is a beginning of a classification of the essential elements and how their interaction causes important concerns. We conclude with a set of annotated example problems from two specific categories that require the user of this paper to demonstrate understanding of critical concepts to achieve the maximum result.

Introduction

During the time since 1965, that serious efforts have been under way to develop strong computer chess programs, a great deal has been learned about the structure of chess. Such things include the fact that brute-force searchers are very tenacious defenders, and that much that humans make a great deal of can be found, essentially en passant by a searching program. The reason that ideas such as board control do not appear in major chess programs is that board control correlates very highly with piece placement. If one side manages to place its pieces well while the other does not, this almost certainly means that the former side dominates the board in such a way as to prevent the opponent from achieving normal good squares for his pieces.

For such reasons the leading programs since 1980 appear to be deficient in knowledge. We understand how depth of search is related to good tactical play. However, a program that does not understand the weakness of a doubled isolated pawn will probably have to search to depths up to 40 ply to discover this fact from more primitive features such as material. Thus, even prodigious searchers require some knowledge [1]. Humans, who search significantly less than computers, require more knowledge. However, some knowledge is required by all those who do not

plan to search to find mates and stalemates. What computer chess has done is to shed considerable light on what the subject of advantages is all about. Certain advantages such as board control are merely correlates of other advantages. Pawn structure advantages are not. Further, there seem to be advantages that depend very much on a great deal of context. This is the subject of study of the "graduate school of chess knowledge" which elevates a player from class "A" to International Class.

Despite the ability of the top few programs to defeat all but the best few hundred players in the U.S., there remains considerable skepticism among artificial intelligence researchers (see e.g. [6, 10] as to whether these results represent competence or performance driven success.

The matches between World Champion Gary Kasparov and Ex-World Champion Anatoly Karpov against DEEP THOUGHT (October 1989, 2-0, and February, 1990, 1-0, respectively, won by the humans) suggest that there is still some way to go before top programs can seriously challenge the World Champion.

Usually performance can be achieved in two different ways:

- By essentially defeating all those rated below and losing to those rated above.
- By achieving non-uniform results against a spectrum of players that averages to a given rating.

Usually, the more erratic a player's performance (class 2 above), the more "promising" he is, and the earlier he will make his upward move by acquiring the necessary remaining skills. Players in class 1 are thought of as unimaginative and able only to wield a tactical cudgel. Thus, the success of top computer chess programs on the Elo rating scale is much more in accord with the former notion than with the latter. However, programs, in defiance of the "pure tactician" mold, have shown the ability to find unique, sound, beautiful, and important ideas by "totally inhuman methods" with relatively little knowledge.

While there is substantial evidence to support the contention that top programs can perform (obtain results) at the strong master level (namely their Elo ratings), there is little evidence to suggest that top programs have a deep

understanding of the profound strategic consequences of a given move -- in this sense programs lack competence, for they will still make moves in a particular position which can be clearly assessed as bad according to elementary or classical heuristics.

We do not wish to sanctify knowledge for its own sake. Knowledge, like anything else in a chess playing entity must be able to pay its own way, and only "useful" knowledge should be acquired. Thus, while researchers into human thinking may find very large amounts of knowledge there, we will attempt to show that computers have an excellent method for distinguishing between useful and superfluous knowledge.

Background and Methods

The thought processes of chess players were studied by de Groot [5] and he determined that Grandmasters look at somewhere between 50 and 200 nodes in the process of deciding on the best move in a given position. Chase and Simon [3, 4] demonstrated that it is the "chunking" of familiar chess specific patterns (groups of pieces) which deems Grandmasters superior to novices in their ability to recall a position from short term memory. They, and later Nievergelt [11] estimated that the number of chess-specific concepts stored in a master's head is somewhere between 50,000 and 100,000. Further work by Kopec and Bratko [9], and Berliner and Campbell [2] gave evidence for the important role played by pawn structures and pawn groups in strong players' abilities to determine what is critical in a chess position. However to this date, we know of no serious effort to develop a comprehensive taxonomy of important chess ideas. It is clear that many important chess themes and ideas have been subsumed by the tactical ability of top programs.

World Champion Kasparov believes that as programs get stronger, so will humans always be able to find new ways to exploit their weaknesses. Considering that he has the highest chess rating ever achieved (2800+), it seems worth taking his point of view seriously. It is clear that in chess there is a hierarchy to the vast number of ideas, heuristics, concepts, and even first order principles such as "In the Opening it is important to develop your pieces." It is also clear that many decisions by domain specialist experts

(grandmasters in chess) do not follow from first principles. Grandmasters, like experts in any domain, know when to break the rules. They know when pieces should be moved more than once in the Opening neglecting the development of other pieces. Another example is the classic advice "Passed pawns must be pushed." However, when does a passed pawn become weak after it is pushed? As general rules for the novice and student, such heuristics are excellent for learning and improving play. There are, nonetheless, further refinements of general heuristics which require special knowledge about, for example, outside passed pawns, protected passed pawns, blockaded passed pawns and passed pawn masses, amongst others. More abstract chess concepts such as king safety, weak square complexes, notions of defense, time, space, and connectivity become harder to quantify, and bring to the fore issues of meta-concepts -- that is, how should concepts and their interaction be quantified?

Earlier attempts at chess test suites ([12, 9]) have either been put together for a different purpose or were not wide-ranging enough to provide a comprehensive test of chess understanding. Results on the Bratko-Kopeck Test showed a strong correlation with ratings, especially for humans. For computer programs, none of which had at the time achieved master ratings, results on the 12 tactical positions were higher than would have been predicted by their ratings, but performance on the 12 lever test positions clearly indicated that this was an area where more domain specific positional knowledge was needed, i.e. tactical depth could not compensate for the lack of this knowledge. Further evidence of the problems presented by lever positions was demonstrated by Marsland [7] from the results of his administering the test to the applicants in the 6th World Computer Chess Championship in Edmonton. The criticism that 24 positions cannot reasonably be a representative sample from the space of the estimated 10^{43} chess positions is certainly valid.

The taxonomy of chess ideas being presented here is only the beginning of a classification process which in some sense can never be complete. Our ultimate goal is to identify as many relevant classes as realistically feasible with examples across all phases of play. Problems in the Opening will be based on "common knowledge" themes derived from the vast

repertoire of human experience. The choice of Middlegame positions is based on difficult, deep or complex tactical themes as well as deep-rooted, critical or refined strategic ideas. Endgame classifications are focused on demonstrating the importance of domain specific knowledge and distinguishing between similar looking, but different positions, or on illustrating the depth belied by simple-looking positions.

Chess programs may be viewed as highly tuned expert systems. Although results from tournament play are quite revealing, an extensive (but not necessarily exhaustive) test suite, such as the one we are constructing, is necessary to identify, in an organized manner, the important concepts that still elude computer chess programs. In their final state these test positions will be available in machine readable form. A taxonomy of positions across the three phases of play, Opening, Middlegame, and Ending, is presented in Table 1 below, with 12 sample positions (6 Ending examples, 3 Opening examples, and 3 Middlegame examples) and their solutions following.

To make the test suite meaningful, we have attempted to produce a uniform scoring system that evaluates both the judgement of the value of the position and finding the correct way to proceed. The points achievable by the user will vary according to the difficulty of the position, and the scoring takes account of things such as alternate ways of proceeding. We conceive of these positions being used by both humans and machines. In both cases, the attempt is being made to identify some facet of chess understanding which may be lacking.

OPENING CATEGORIES (O)

01. Gambits Accepted or Declined
02. Return Material at correct time
03. Pawn for Development (e.g. Poison Pawn Variation, Najdorf)
04. Bishop Outside/Inside pawn chain
05. Multiple piece moves to advantage
06. Good and Bad (attackable) Centers
07. When to (and not to) fianchetto
08. The Options Principle
09. Critical pawn moves for space.
010. Bishops for Knights (examples)
011. Choice of where to develop B/N.
012. Avoiding "shutout" of a piece(s)

- O13. Qualitative vs. Quantitative Development
- O14. Loss of Time due to multiple piece or pawn moves

MIDDLEGAME CATEGORIES (M)

Pawn Play (MP)

- MP1. Pawn advances with idea of opening files (involving pawn exchanges)
- MP2. Minority Attack
- MP3. Preserving Pawns on side where win occurs (e.g. Outside P.P.; Potential Outside P.P)
- MP4. Pawns on d5 or e5; center pawns blockaded or not.
- MP5. Structures related to minor pieces.
- MP6. Decisive Pawn weakenings.
- MP7. Advances in front of king.
- MP8. Advancing for space.
- MP9. Keeping or relinquishing tension.
- MP10. Attacking with pawns
- MP11. Central pawn mass.
- MP12. Creating pawn structural strengths/weaknesses.
- MP13. Maturing attacks.
- MP14. Pawn Structures
 - (a) Doubled Pawns
 - (b) Meaningful Majorities (Handling)
 - (c) Isolated
 - (d) Backward
 - (e) Passed
 - (f) Hanging
 - (g) Chains

Piece Play (Positional) (MPPP)

- MPPP 1. Normal material values do not apply
 - MPPP 1a. Exchange Sacs
 - MPPP 1b. Queen Sacs
- MPPP 2. Piece Sacrifices
- MPPP 3. When to give up the bishop pair; which one, how.
- MPPP 4. Superior Knight
- MPPP 5. Position of Rooks
- MPPP 6. Choice of Rook Placement (where options exist)
- MPPP 7. Which side to castle on (if at all)

- MPPP 8. Piece Regroupings (R,B,N,Q)
- MPPP 9. N on rim/ good/bad
- MPPP 10. Trapped Pieces
- MPPP 11. Freeing Trapped Pieces
- MPPP 12. Bishops of opposite colors;
 - 12a. Attacking with
 - 12b. Exchanging queens with
- MPPP 13. King walk
- MPPP 14. Removing the key defender of an opponent's position

Piece Play (Tactical) (MPPT)

- MPPT 1. Desperado
- MPPT 2. Double Attack
- MPPT 3. Hanging Pieces
- MPPT 4. Interference
- MPPT 5. Decoy
- MPPT 6. Overload
- MPPT 7. Skewer
- MPPT 8. Trapping
- MPPT 9. Pin

ENDINGS (E)

E1: R + P Endings (RPE)

- RPE 1. Rooks Threatening Pawns
- RPE 2. Rooks behind, on side of, in front of P.P.
- RPE 3. Connected P.P.
- RPE 4. Position of defending king
- RPE 5. Superior Pawn Structure

E2: B vs. N (superior/weak)

E2a: B vs. B Same color, bad Bishop.

E3: Complex Pawn Endings

E4: Q + P endings

E5: R + B vs. R + N

E6: R + P's vs. Two Minor Pieces

E7: R + B + B + extra pawn(s) vs. R + R + N

E8: Mating attacks in endings

E9: General: Not how many pawns, but their position counts

E10: Exchange up with pawns on board

E11: Elementary Endings

(a) R + P vs. R

(b) R vs. N

(c) R vs. B

(d) Q vs. R

(e) Q vs. BB

(f) BB vs. N

(g) others

Table 1: A TAXONOMY OF CHESS IDEAS

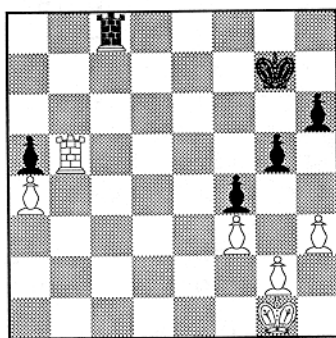


Figure 1
Black to Play

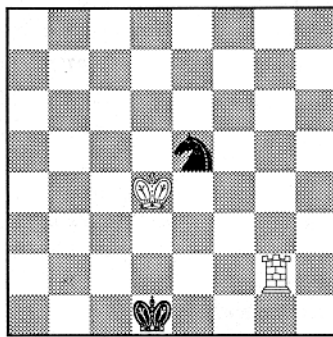


Figure 2
Black to Play

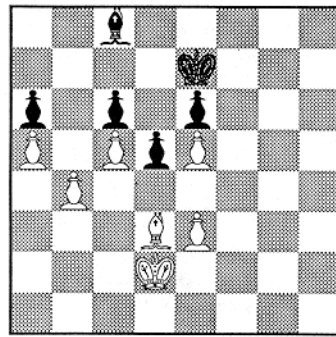


Figure 3
White to Play

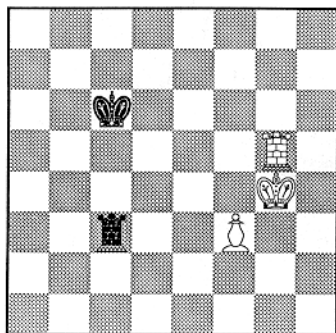


Figure 4
White to Play

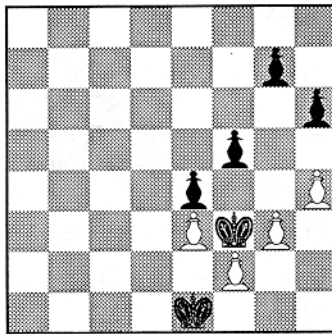


Figure 5
Black to Play

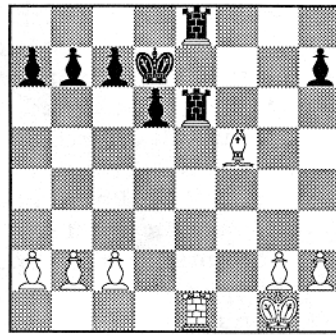


Figure 6
White to Play

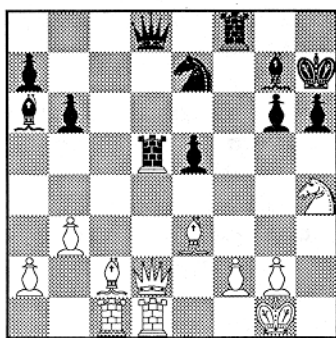


Figure 7
White to Play

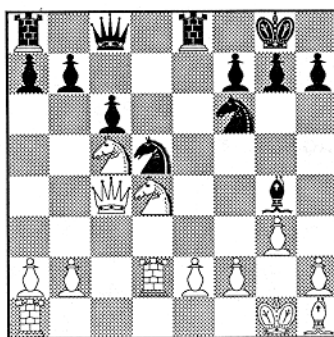


Figure 8
Black to Play

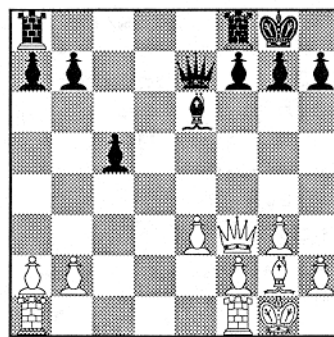


Figure 9
White to Play

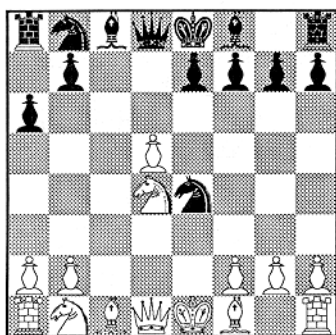


Figure 10
White to Play

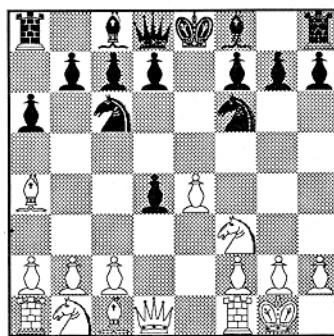


Figure 11
Black to Play

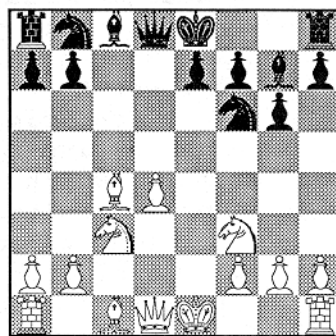


Figure 12
White to Play

Solutions to Positions

The five position categories are as follows:

1. White is winning
2. White is superior
3. The position is equal
4. Black is superior
5. Black is winning.

Figure 1: White is threatening the Black a-pawn. The correct move for Black is to prepare for the resulting ending by finding an active position for his rook with 1. ...Rc2 or better still 1. ...Rc1+. 1. ...Ra8? is the wrong move when Black is likely to lose after a continuation like 2.h4! Kg6 3. Kh2 Ra7 4. Kh3 Kh5 5. hxg5 hxg5 6. g3! fxg3 7. f4. The guiding principle for rook and pawn endings is that the defending side's rook must remain active. After 1. ...Rc1+ 2. Kh2 Ra1 3. Rxa5 Kg6 4. h4 gxh4 5. Kh3 h5! 6. Kh2 h3! 7.gxh3 Ra2+ Black draws easily. It is positions like this which distinguish between good heuristics like "Rooks belong behind passed pawns" and knowing when and how such knowledge is critical. Category 3. Difficulty points = 4

Figure 2: The only drawing move is 1. ...Nc6+ and then after 2.Kd5 Nb4+ and 3. ...Nc2, when the BK and BN are safely together. If 1. ...Nf3+? 2. Ke3 Ne5 (2. ...Ne1 3.Rg1) 3.Rg5! (the only winning move) Nc4+ 4.Kd3 and White wins the N in some 10 moves. This is one of the first positions found to be an exception to the heuristic for this ending "Keep K and N together at all times". What this shows is that the superficial application of the rule does not always work; i. e. keeping the knight and king together means not allowing them to be driven apart in the near future. This position also sprouted many others in this ending which were exceptional and illustrative of the depth and beauty which can underly this simple-looking ending. Category 3. Difficulty points = 4.

Figure 3: Black has a very bad bishop, but the only way this can be exploited is through the manoeuvre 1.Kc3 2.Kd4 and only then e4 opening lines decisively. If instead 1.e4? d4! and after White wins the P/d4 he will find that there is no way to enter Black's position. With the recommended method, either the white bishop will get a significant role (after 3. e4 dxe4 4. Bxe4) or

white will get another passed pawn after 3. e4 and 4. exd5. This position is clearly of the long strategical kind whereby White must accomplish small goals in order to achieve the major goal of winning. Category 1. Difficulty points = 4.

Figure 4: In this position it is critical that White can cutoff the BK in order to prevent him from reaching the vicinity of the queening square. The winning move is 1.Re5. The continuation in a game won by Tal went: 1. ...Kd6 2. Re8 Kd7 3. Re1 Rc8 4. f4 Rg8+ 5. Kf5 Rf8+ 6. Kg5 (1-0). Cutting of the defensive king to facilitate reaching the "Lucena Position" (whereby the WK controls the queening square) is fundamental to such endings. Category 1. Difficulty points = 2.

Figure 5: Black obtains an outside passed pawn via the breakthrough 1. ...f4!. If 2. exf4 e3 leads to the win of all of White's pawns. Or if 2.gxf4 Kg4 and Black obtains the winning outside passed pawn. A critical continuation is: 3.Ke2 h5 4.f5 Kxh4 5.f3 gxf3+ 6.Kxf3 Kg5 7.e4 h4 and Black wins. This example demonstrates the importance of king domination and outside passed pawns. Category 5. Difficulty points = 3 for correct category; 3 for move f4; 3 more if you could demonstrate the win in all variations.

Figure 6: After 1.c4! there follows intricate pawn play based on the theme of maintaining a permanent pin on the R/e6 whereby White will only simplify when the K+P ending is trivially won. If Black plays 1. ...a5 (to prevent b4) then White replies b3, then a3, and b4, enabling ...d5 to be answered by c5 when the pin remains. Black's other tries are to play 1. ... c6 or b6, when White now plays 2. b4!, and any further d5 black can still be answered by c5, maintaining the stranglehold on Black. Category 1. Difficulty points = 4.

Figure 7: White has just lost a pawn on d5 but gets a winning attack after 1. Qxd5! demonstrating that the BN/e7 is overworked, i.e. Nxd5 2.Nxg6 threatening both 3. Rxd5 and 3. Nxf8+ to which B has no defense. White quickly recovers more than enough material for the Q and a continuing attack against the exposed BK. The combination has served to expose the weaknesses around the BK, especially on the b1-h7 diagonal. Category = 1. Difficulty points = 3 for correct category; 3 for convincing variations.

Figure 8: A very famous position in which the seeds of Alekhine's (B) attack against Reti (W) began with 1. ...h5, which was followed by ... h4,

resulting in a weakening of the fianchetto pawn fortress in front of the WK, which culminated in a brilliant attack focused on softened g3 square. Category = 4. Difficulty points = 2 for correct category; 3 for correct move.

Figure 9: White should try to get his K-side pawn majority mobilized with 1.e4 followed by Qe3 and f4 and then try to generate some counterplay against the BK. If the position reaches an ending (as occurred in Marshall-Capablanca, 1909) Black can expect to win due to his Q-side pawn majority. Category = 4. Difficulty points = 2 for category; 3 for correct move.

Figure 10: This is a position from opening in the famous Alekhine-Wolfe, Pistyan, 1922, game. It is a unique example of when to break a principle. 1. d4 d5 2. Nf3 c5 3. c4 cxd4 4. cxd5 Nf6 5. Nxd4 a6?! 6. e4!! Nxe4. Now 7. Qa4+! Bd7 8. Qb3 Nc5 8. Qe3! g6?! 9. Nf3 Qc7 and now 10. Qc3! The multiple Q and N moves have kept Black from completing his development, and now he must decide between a permanent weakness by f6 or renouncing castling by Rg8. Clearly, the rule, "Do not move a piece more than once in the opening" was one which Alekhine knew exactly when to break. Category = 1. Difficulty points = 3 for move and category.

Figure 11: Black's sixth move and already it is a critical one in determining the course of play. Most solid is 6. ...Be7. Other moves, especially 6. ...Nxe4? can quickly lead to serious problems for Black. In certain sharp opening positions you simply must know the moves. Category = 1. Difficulty points = 3 for move and category.

Figure 12: Hitech-Denker, 1988. 1. e4 c5 2. c3 d5 3. exd5 Qxd5 4. d4 Nf6 5. Nf3 cxd4 6. cxd4 g6? 7. Nc3 Qd8 8. Bc4 Bg7 and now what is the right move? 9. Qa4+ must be answered by Kf8 as the played Nbd7 allows 10. Bxf7+ Kxf7 11. Ng5+ Ke8 12. Ne6 Qb6 13. Qc4! winning and 10. ... Bd7 11. Qb3 O-O 12. Ne5 wins a pawn smoothly. Category = 1. Difficulty points = 3 for move and category.

Acknowledgements

The authors wish to thank Robert A. Cicogna and A. John Ostuni for their technical assistance.

References

- [1] Berliner, H. J.
Search vs. Knowledge: An Analysis from the Domain of Games.
In Elithorn and Banerji (editor), *Artificial and Human Intelligence*. Elsevier Science Publishers, 1984.
- [2] Berliner, H., and Campbell, M.
Using Chunking to Play Chess Pawn Endgames.
Artificial Intelligence 23(1), May, 1984.
- [3] Chase, W. G. and Simon, H. A.
Perception in Chess.
Cognitive Psychology 4(1), January, 1973.
- [4] Chase, W. G. and Simon, H. A.
The Minds Eye in Chess.
In Chase, W. G. (editor), *8th Annual Psychology Symposium Volume: Visual Information Processing*, chapter 8, pages 215-281. Academic Press, 1973.
- [5] de Groot, A.
Thought and Choice in Chess. Mouton, The Hague, 1965.
- [6] Donskoy, M. and Schaeffer, J.
Perspectives on Falling from Grace.
In Marsland, T. A. and Schaeffer, J. (editor), *Computers, Chess and Cognition*. Springer-Verlag, New York, 1990.
- [7] Marsland, T. A.
The Bratko-Kopec Test Revisited.
In *Proceedings of the Workshop on New Directions in Game-Tree Search*, pages 135-139. Edmonton, Alberta, Canada, May, 1989.
- [8] Hort, V., and Jansa, V.
The Best Move. RHM Press, New York, 1980.
- [9] Kopec, D. and Bratko, I.
The Bratko-Kopec Experiment: a comparison of human and computer performance in chess.
In Clarke, M.R.B. (editor), *Advances in Computer Chess* 3, pages 57-72. Pergamon Press, 1982.

- [10] Kopec, D.
Advances in Man-Machine Chess Play.
In Marsland, T. A. and Schaeffer,
J. (editor), *Computers, Chess and
Cognition*. Springer-Verlag, New York,
1990.
- [11] Nievergelt, J.
Information content of chess positions:
implications for chess-specific
knowledge of chessplayers.
SIGART Newsletter (62):13-15, 1977.
- [12] Reinfeld, F.
Win At Chess.
Dover Books, New York, 1958.