# Search Versus Knowledge in Human Problem Solving: A Case Study in Chess

**Ivan Bratko, Dayana Hristova and Matej Guid**

**Abstract** This paper contributes to the understanding of human problem solving involved in mental tasks that require exploration among alternatives. Examples of such tasks are theorem proving and classical games like chess. De Groot's largely used model of chess players' thinking conceptually consists of two stages: (1) detection of general possibilities, or "motifs", that indicate promising ideas the player may try to explore in a given chess position, and (2) calculation of concrete chess variations to establish whether any of the motifs can indeed be exploited to win the game. Strong chess players have to master both of these two components of chess problem solving skill. The first component reflects the player's chess-specific knowledge, whereas the second applies more generally in game playing and other combinatorial problems. In this paper, we studied experimentally the relative importance of the two components of problem solving skill in tactical chess problems. A possibly surprising conclusion of our experiments is that for our type of chess problems, and players over a rather large range of chess strength, it is the calculating ability, rather than chess-specific pattern-based knowledge, that better discriminates among the players regarding their success. We also formulated De Groot's model as a Causal Bayesian Network and set the probabilities in the network according to our experimental results.

## 1 Introduction

Consider human solving of mental tasks such as theorem proving, symbolic manipulation problems, and classical games such as chess or checkers. A general, widely accepted computational model of solving such problems involves searching among alternatives (Newell and Simon 1972). In games, for example, this amounts

I. Bratko (✉) · M. Guid
Faculty of Computer and Information Science, University of Ljubljana,
Večna pot 113, 1000 Ljubljana, Slovenia
e-mail: bratko@fri.uni-lj.si

D. Hristova
University of Vienna, Vienna, Austria



L. Magnani and C. Casadio (eds.), *Model-Based Reasoning in Science and Technology*, Studies in Applied Philosophy, Epistemology and Rational Ethics 27, DOI 10.1007/978-3-319-38983-7_32

to searching through "my" possible moves, then for each "my" move considering all possible opponent's replies, then again considering my possible moves for each of the opponent's replies, etc. This search may stop when positions are encountered that are estimated without doubt as drawn, or good or bad for one of the sides. The result of search is "my" move (if exists) that guarantees my win against any possible opponent's reply.

In fact, this simple algorithm is *in principle* sufficient to solve problems of the types mentioned above. However, in cases of high combinatorial complexity of problems, such as chess where there are many alternatives at each step, this is practically infeasible because it takes too long for humans, and even for computers. Therefore the search has to be carried out intelligently, relying on search heuristics that are based on the problem solver's knowledge about the domain. Newell coined the term knowledge search (1990) describing the way an agent uses their *directly available* long-term knowledge to bear on the current situation in order to control the search. These heuristics guide the search in promising directions and thus reduce the complexity of search needed to solve problems. The usual, empirically observed relation between the amount of solver's problem-specific knowledge and the amount of search required is: the more knowledge the problem-solver possesses, the less search is needed.

In this paper, we study this trade-off between knowledge and search in human game playing, and investigate their relative importance for success. To this end, we conducted experiments in human problem solving in the game of chess. The problem for the participants was: given a chess position, find the best move for the side to move. The human problem solving model for this case, relevant to our study, was stated in the classical work by De Groot (1946, 1978) and was also used by Tikhomirov and Poznyanskaya (1966).

De Groot's model of chess players' thinking about best moves conceptually consists of two stages: (1) position investigation, and (2) investigation of possibilities, or search. Stage 1, "position investigation" consists of identifying general properties of a position like "Black king is not well protected, so a direct attack on Black king should be considered", and familiar patterns like "White knight is pinned by the Black bishop, so the knight might be in danger". Such patterns make the player suspect that there might be a way to checkmate Black king, and that White knight might be attacked by other enemy pieces and eventually lost because a pinned knight cannot escape from the attack. However, no search among concrete moves is done at Stage 1. This is done at Stage 2, "investigation of possibilities". This consists of the calculation of concrete moves and variations that may lead to the actual exploitation of the spotted motifs. That is, in our examples, to force the checkmate of the king, or force the capture of the knight. Stage 2 is similar to the usual computational problem solving model that involves search described earlier. In this paper, Stage 1 will be referred to as "detection of motifs", and Stage 2 as "calculation of variations". Strong chess players have to master both components of problem-solving skill: detection of motifs, and calculation of variations. The first component is based on the player's chess-specific knowledge, which reflects the player's general understanding of the game. The player has acquired such an understanding through experience and her study of chess literature. The second component, i.e. calculation, seems to be less specific to chess, and is reflected in the player's ability to reliably calculate in his or her mind the possible variations.

The research question explored in this paper is: which of these two components better separates successful players from unsuccessful ones. By successful players we mean those who find best moves more often.

## 2 Examples of Motifs and Calculation

To further clarify the two stage model of chess problem solving, we now present concrete examples of motifs and calculations.

Figure 1, diagram (a), illustrates "a pin", one of the most common and effective motifs (patterns) in chess. White rook attacks Black knight. The knight cannot escape because this would leave Black king under attack. Black cannot do anything to prevent the loss of the knight, and White wins. Diagram (b) shows how such a motif enables the player to find the right move for White almost without any calculation. The motif itself immediately suggests the winning move rook b1-b7. It then remains to calculate that after all possible Black king moves, White rook can capture the knight. This amounts to searching some 8 positions altogether.

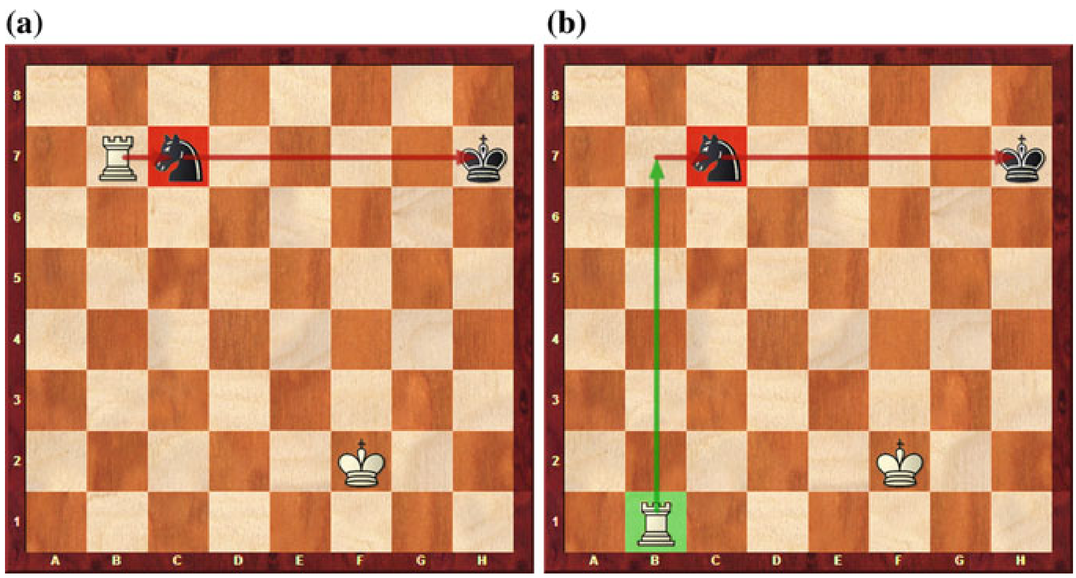


**Fig. 1** Diagram **a**. *Black* to move: One of the simplest and most common motifs in chess, called a "pin". *White rook* is attacking the *Black knight*. The knight cannot escape from the rook's attack because if the knight moves away *Black king* will come under the rook's attack (indicated by the *red arrow*) which means a knight's move is illegal. Therefore the Black knight is said to be pinned, and in fact lost in our case. Diagram **b**, *White* to move: The White player in a glance notices an instance of the familiar pin motif (*Black knight* and *Black king* are both on the same line), so the player will immediately perform a little calculation to see whether this motif can actually be exploited. The winning move is: rook moves from square b1 to b7 (*green arrow*) to pin the knight. After any Black's reply, the knight will be captured by the rook

Without the motif, the unguided calculation would have to consider incomparably more positions, something like 30.000. This estimate can be roughly worked out as follows. Let White perform, for example, the breadth-first search. To solve the problem of Fig. 1b, the required depth of search is four half-moves. That is: first level moves by White, then Black's replies, and then the second level moves by White and another level of Black's replies. In each position, White has roughly 22 possible moves (maximally 14 moves by the rook plus 8 moves by the king). Black has roughly 16 possible moves in each position (maximally 8 moves by the knight and 8 by the king). Taking into account gradual reduction of the number of moves with increasing level, this means something roughly in the order of 30.000 positions; compared with 8 when the motif is available.

It should be noted that the pin is a very general concept. It does not have to necessarily involve a rook and a knight and a king, as in our example. A pin occurs between any three pieces where the pinning piece is a long range piece that moves either horizontally, vertically or diagonally (a queen, a rook or a bishop), the pinned piece and the target piece that is indirectly threatened by the pinning piece:

Pinning piece ——→ Pinned piece - - - - - -> Target piece

The arrows show the line of attack by the pinning piece. The pinned and the target piece are of the same colour, and the pinning piece is of the opposite colour. The dashed line only becomes available to the pinning piece when the pinned piece moves away and frees the path to the target piece. Typically, a pin is all the more effective if the target piece is more valuable than the pinned piece. If the pinned

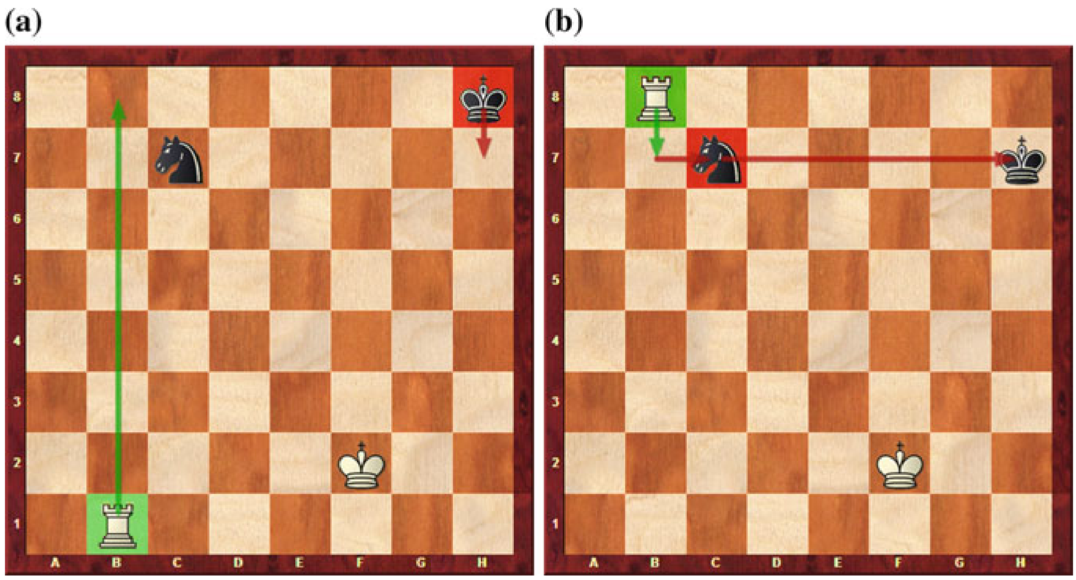


**Fig. 2** Diagram **a**: There is no pin yet on the board, but can *White* achieve a pin by forcing the *Black* pieces into the same line? Yes, with *White rook* move from b1 to b8 (*green arrow*). After the move *Black king* is in check and has to move to the 7-th rank, from h8 to h7 or g7. Then the familiar motif appears (diagram b)

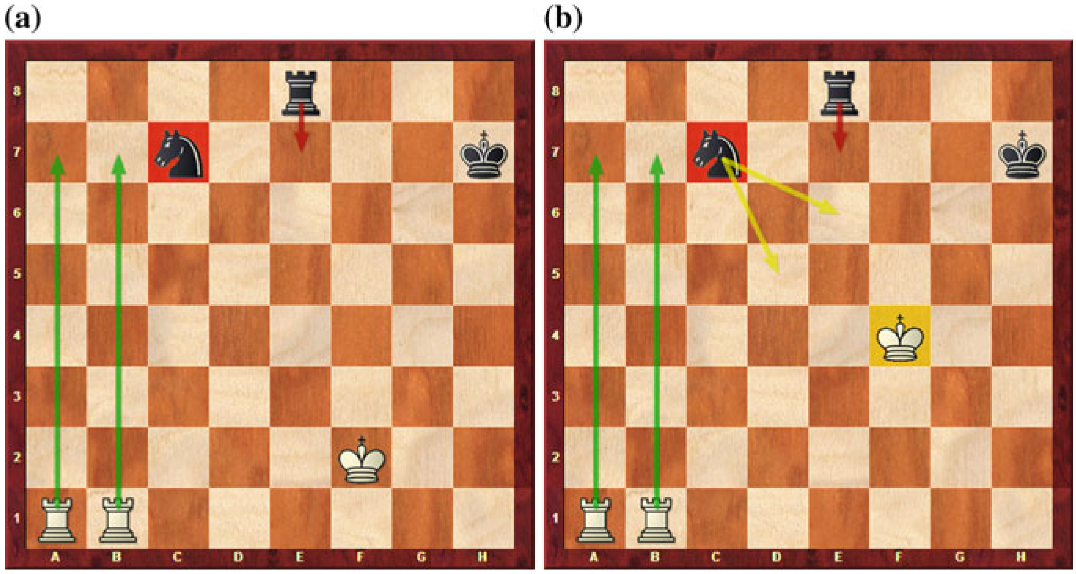


**Fig. 3** Diagram **a**: The pin motif in a more complex position with an extra rook for *White* and *Black*. The same idea works for *White* as in Fig. 1, but the calculation here is more complex. White may start with rook from b1 to b7, pinning the knight. Now *Black* has more defensive resources enabled by the presence of *Black rook*. One way is to play Re8-e7, defending the knight with the rook, and interrupting *Black knight* pin against *Black king*. However, the knight is still pinned, this time against *Black rook*. So *White*, seeing this, logically increases the pressure on *Black knight* by playing Ra1-a7, further exploiting the pin. *Black knight* may try to flee to d5, also defending *Black rook*. But the rook is then attacked by two *White rooks*, so White captures twice on e7 and wins. Instead of moving knight to d5, *Black* may try to check *White king* with rook f7 check. Now *White* has to be careful and move the king to e2 which wins eventually, but not to g2 (or g1 or g3) because after that, *Black* can deliver another check by rook f7-g7. After *White king* g2-h2 *Black* has no further checks, but can save himself with Nc7-e6. Suddenly, *Black* pieces have reorganised themselves, *Black rook* is now protected by both the king and knight, and position is drawn. This example nicely shows the role of calculation. The calculation is driven by the motifs, but the final truth is determined by calculation. Diagram **b**: This is the same as diagram (**a**) except that *White king* is now at f4 instead of f2. This small difference offers *Black* additional defensive possibilities and further complicates the calculation drastically. After *White rooks* have doubled on the 7th rank, *Black knight* can move out of attack to d5 or e6 with check. This way *Black* gets out of the pin trouble. But in the process, the knight gets misplaced and cannot return towards his king. Eventually White pieces can trap the knight and win. We do not give concrete variations because they become numerous and very long

piece moves away to escape the attack, the pinning piece may capture the target, possibly winning even more material than just the pinned piece.

Figure 2 shows an example where a pin cannot be immediately created in one move, but can be forced in two moves. The calculation now becomes more demanding, but not significantly so because the motif still guides search very effectively. Any reasonably strong player will immediately look for a way of forcing a pin, and find it in no time.

Figure 3, diagram (a), shows a rather more complex example on the same theme. The pin motif is the same, but two additional rooks now significantly complicate the calculation. But even so, the calculation proceeds in more or less a similar way as in the previous examples according to the familiar mechanisms of exploiting a pin.

Estimates of the number of positions to be searched are: (1) a few tens if guided by the motif, and (2) some 6.000.000 if unguided, which is completely outside human calculation capability. This example illustrates the typical law that the difference between the complexity of guided vs. unguided search grows exponentially with the required search depth. Figure 3b is a small variation of 3a in which White king is at f4 instead of f2. This example shows that, although the main pin motif remains the same as in 3a, the small variation enormously increases the calculation complexity and even brings in several additional motifs in different variations.

## 3 Experimental Setup

We investigated our research question with an experiment in which 12 chess players of various chess strengths were asked to solve 12 tactical chess problems. A chess position is said to be *tactical* if finding the best move requires calculation of variations, in addition to detecting tactical motifs in the positions.

Our players' chess strength, in terms of official FIDE chess ratings was in the range between 1845 and 2279 rating points. The strength of registered chess players is officially computed by FIDE (World Chess Federation), using the Elo rating system. This rating system was designed by Arpad Elo (Elo 1978). For each player, this rating is calculated and regularly updated according to the players' tournament results. The rating range of our players, between 1845 and 2279, means that there were large differences in chess strength between the players. The lowest end of this range corresponds to club players, and the highest end to chess masters (to obtain the FIDE master title, the player's rating has to reach at least 2300 at some point in their career). There were actually two chess masters among our participants. According to the definition of the Elo rating system, the expected result in a match between our top ranked player against our lowest ranked player would be about 92 % against 8 % (the stronger player winning 92 % of all possible points).

In addition to the differences in chess strength expressed through chess ratings, one could also consider other differences between the players. One such factor might be the chess school where a player was taught, or the particular instructor by whom the player was trained. However, in this paper we did not explore the effects of such additional factors. The 12 chess problems were selected from the Chess Tempo web-site (www.chesstempo.com) where the problems are rated according to their difficulty. Chess problems are rated in a similar way as the players. However, in Chess Tempo a problem's rating is determined by the success of the players when solving the problem. The principle is as follows: If a weak player solved a problem then this counts as strong evidence that the problem is easy. So the problem's rating goes down. If a stronger player solved the problem, the problem's rating still goes down, but not as much as for a weak player. On the contrary, if a strong player failed to solve the problem, this counts as strong evidence that the problem is hard, and the problem's rating increases. In detail, a problem's rating in ChessTempo is determined by using the Glicko rating system (Glickman 1999) which is similar to the Elo system. The Glicko system, at difference with Elo, takes into account the time a player has been inactive. In cases of longer inactivity, the player's rating becomes uncertain. To illustrate the meaning of ratings in ChessTempo, a player with rating 2000 has a 50 % chance of correctly solving a problem with the same rating of 2000. The same player has a 76 % chance to solve a problem with the rating 1800, and a 24 % chance to solve a problem rated 2200.

In our selection we ensured a mix of problems that largely differ in their difficulty. Our selected positions were all tactical chess problems, randomly selected from Chess Tempo according to their difficulty ratings. Based on their Chess Tempo ratings, our problems can be divided into three classes of difficulty: "easy" (2 problems; their average Chess Tempo rating was 1493.9), "medium" (4 problems; average rating 1878.8), and "hard" (6 problems; average rating 2243.5). While the problems within the same difficulty class have very similar difficulty rating, each of the three classes is separated from their adjacent classes by at least 350 Chess Tempo rating points. Some problems have more than one correct solution. To ensure correctness, all the solutions were verified by a chess playing program.

The experimental setup was as follows. Chess problems, that is chess positions in which the participant was asked to find a winning move, were displayed as chess diagrams on a monitor, and the players' solution moves were recorded. Allowed solving time per position was limited to 3 min.

During the player's problem solving, the player's eye movements were tracked by an eye-tracking device EyeLink 1000 and recorded into a database. In the last decades, with wide availability of eye tracking devices, studying eye movements has turned into one of the main methods of research in chess decision making (Reingold and Charness 2005). The processing of recorded eye-movements reveals roughly on which squares of the chessboard the participant was focussing at any time during problem solving.

After the player finished with the 12 problems, a retrospection interview was conducted in which the player described how he or she approached the problem. It was possible to detect from these retrospections what motifs were considered by the player, and roughly how the calculation of variations driven by the motifs was done. Other details of the experiments are described in (Hristova et al. 2014a, b) where the question of automated assessment of the difficulty of chess problems was tackled.

In this paper, we analyse the experimental data with respect to the research question of this paper. The relevant experimental data includes the following. For each player and position, the relevant information consists of: (1) correctness of the solution proposed by the player, (2) the motifs considered by the player in comparison with the motifs needed to solve the problem, and (3) the correctness of the calculation of variations. The motifs considered were found through the players' retrospections, and to some extent verified by the eye movement data, although this verification cannot be done completely reliably.

Figure 4 shows an example of how the eye tracking data can be used. It should also be noted that not all relevant motifs in a position were needed to solve the

(a) (b)

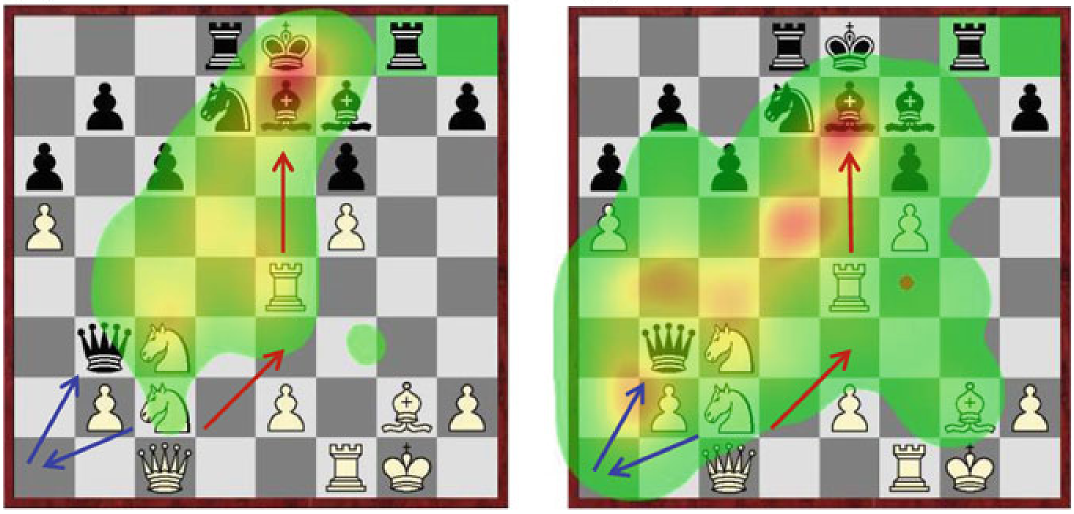

**Fig. 4** One of the more difficult test positions. The colouring was extracted from eye tracking data, and shows the intensity of eye focussing on areas of the chess board by a player (*red* is the most intensive, then follow *yellow* and *green*; *white* means no significant focussing). The two diagrams correspond to two players, call them A and B. There are two main motifs in this position: (1) attack *Black king* who is placed very uncomfortably and can be attacked e.g. by the *White rook* at square e4 (*middle* of the board) and *White Queen* moving to square e3. The directions of attack by these moves are indicate by *red arrows*; (2) *Black Queen* is surrounded by *White* pieces and it looks that it might be trapped; some relevant moves by a *White Knight* against *Black Queen* are shown with blue arrows. Calculation of concrete variations reveals that attacking *Black King* does not win for *White*, but trapping *Black Queen* does (by moving the Knight from c2 into the bottom left corner of the board). Player A, according to retrospection, considered both motifs and solved the problem. Player B only considered attacking *Black King* and failed to solve the problem. This difference in motif detection by the two players is indicated in the eye tracking images, although rather subtly

problem. Some of the motifs did not give rise to a winning move. Sometimes there were several alternative winning moves, and there were accordingly several alternative motifs, any of them being sufficient to find at least one of the solutions. Some winning moves were derived from a calculation that required a combination of more than one motif.

These possibilities had to be taken into account when deciding whether the player detected a complete set of motifs needed to carry out correct calculation. To this end, for each position and each possible solution of the position, we defined the "standard" set of motifs necessary and sufficient to find the solution. In defining the standard sets of relevant motifs, we took into account all the motifs mentioned by all the players. In very rare cases when needed, we had to add motifs that fully enabled correct calculation for each possible solution. In doing this, we used our own chess expertise (two of us have chess ratings over 2300 and 2200 respectively). As we verified all the solutions and corresponding chess variations by a chess program, we believe that there would be very little room for defining reasonable alternative standard sets of motifs.

# 4 Results

## 4.1 Experimental Data and Basic Statistics

Our experimental data consisted of the performance of 12 players in all 12 positions, that is 144 players' solutions. Out of these, we omitted four cases that were meaningless because the player misunderstood the task (confused the side to move in the position). This gives a total of 140 proposed solutions that we analysed in the sequel (144 − 4 = 140).

Here are some results relevant to our research questions. The players correctly found a winning move in 89 out of 140 problems, i.e. 63.6 %. The players correctly detected all the motifs relevant to a solution in 121 problems (86.43 %). Note that this is the percentage of cases in which the players perfectly detected relevant motifs. The calculations were completely correct in 35.7 % of the problems. It should be noted that a problem was often solved correctly even in the case of imperfect calculation or imperfect detection of motifs.

In the sequel we slightly refine the possible outcomes and introduce the following notation:

- Variable S stands for "success", that is correct solution found.
- Variable M stands for the event "motifs detected". For a problem and a player, M is true if the player correctly detected all the motifs in the problem position that are relevant to a solution of the problem (perfect detection of motifs); otherwise M is false.
- Variable C stands for "calculation correctness". For a problem and a player, C may take one of three possible values:
  - C = CC if the player's calculation in the position is completely correct; that is, it clearly states the critical variations
  - C = CA if the player's calculation is "adequate". That is, the calculation is basically correct, it does suggest a correct move to play (although possibly with a bit of luck), but it is incomplete and/or indicates the player's uncertainty. For example, the calculation is accompanied by the player's comments like "I was not able to calculate everything", "I had to rely on intuition", etc.
  - C = CI if the player's calculation is clearly incorrect, although it may, through sheer luck, even suggest a correct move to play, but for wrong reasons

CC (calculation correct) occurred in 35.7 % of all 140 cases, CA (calculation adequate) occurred in 22.1 % of all cases, and CI in 42.1 %. Note that these do not exactly sum up to 100 % because of rounding errors. Taking these relative frequencies as simple estimates of probabilities we have:

P(CC) = 0.357
P(CA) = 0.221
P(CC v CA) = 0.579
P(CI) = 0.421

Again, the numbers above do not sum up exactly due to rounding errors.

The following results are indicative of the importance of motif detection. Relevant motifs were not perfectly detected in 19 problems. In none of these problems, the calculation was correct; moreover, it was not even adequate in any of them. Estimating probabilities by relative frequencies, we have:

P(CC | ~ M) = 0/19 = 0.0
P(CC v CA | ~ M) = 0.0

From this, one may conclude that it is very unlikely to perform correct or adequate calculation without relevant motifs.

On the other hand, a successful solution may be found with a bit of luck even in the case of incorrect calculation. This happened in 8 out of 59 cases, so relative frequency is:

P(S | CI) = 0.1356

Even more, a successful solution may be found by luck in the absence of detected motifs and under incorrect calculation. This happened in 2 out of 19 cases, giving relative frequency:

P(S | ~ M & CI) = 0.1053

Regarding our question about the relative importance for success between detection of relevant motifs and calculation, correlations between some of the variables in our domain are important. These variables are: success in a player finding a correct move in a given position, detection of motifs, and correctness of calculation. We computed Pearson's sample correlation coefficients, which requires numerical input data. To this end we defined corresponding numerical variables as follows:

- Success = 1 if the problem was successfully solved; otherwise Success = 0
- MotifsDetected = 1 if all the relevant motifs were detected in the position, otherwise MotifsDetected = 0
- CalculationOK = 1 if the calculation was correct or adequate (in our notation CC or CA), otherwise CalculationOK = 0 (i.e. calculation incorrect, CI)

The correlations between pairs of these variables are:

r(Success, Motifs-Detected) = 0.4368
r(Success, CalculationOK) = 0.8870
r(MotifsDetected, CalculationOK) = 0.4643

It should be noted that the correlation between success and correctness of calculation is much higher than the other two correlations above. This confirms that the correctness of calculation is a more important predictor of success than the detection of motifs. This can be explained by the facts that motifs are highly needed for correct calculation, but they are not highly sufficient for correct calculation.

## 4.2 Causal Bayesian Network Model of Problem Solving in Our Domain

We can formulate a probabilistic model of our chess problem domain as a Causal Bayesian Network. We will use the three binary variables Success, MotifsDetected and CalculationOK. However, in a Bayesian network it will be more convenient to treat them as Boolean variables rather than numerical, so the numerical value 1 will be replaced by true, and 0 by false.

It is most natural to view the causal dependences between these three events according to the problem solving process: first, the player looks for relevant motifs, then she uses these motifs to drive the calculation which results in a successful or unsuccessful solution. This corresponds to the structure of the Bayesian network in Fig. 5.

Note that the link between MotifsDetected and Success cannot be ignored because Success also depends probabilistically on MotifsDetected when CalculationOK is known.

The probabilities for this network are:

P(MotifsDetected) = 0.8643
P(CalculationOK | MotifsDetected) = 0.6694
P(CalculationOK | ~ MotifsDetected) = 0.000
P(Success | CalculationOK ^ MotifsDetected) = 1.000
P(Success | CalculationOK ^ ~ MotifsDetected) = 1.000
P(Success | ~ CalculationOK ^ MotifsDetected) = 0.150
P(Success | ~ CalculationOK ^ ~ MotifsDetected) = 0.1053

It is interesting to check what happens if we omit the link between MotifsDetected and Success, which makes the structure of the model simpler, more intuitive, and better reflect De Groot's basic model of chess thinking (the stage of motif detection is followed by the calculation of variations, then the solver's solution emerges). The structure then becomes:

MotifsDetected ⟶ CalculationOK ⟶ Success

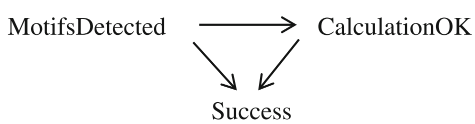


**Fig. 5** The structure of a Bayesian network model of chess problem solving

**Table 1** M, C, S stand for MotifsDetected, CalculationOK, and Success respectively. $P_{full}$ and $P_{simple}$ are the probabilities of the indicated combined events according to the full Bayes model and the simplified model, respectively

| M | C | S | $P_{full}$ | $P_{simple}$ |
|---|---|---|---|---|
| t | t | t | 0.5786 | 0.5786 |
| t | t | f | 0.0000 | 0.0000 |
| t | f | t | 0.0429 | 0.0387 |
| t | f | f | 0.2429 | 0.2470 |
| f | f | t | 0.0000 | 0.0000 |
| f | t | f | 0.0000 | 0.0000 |
| f | f | t | 0.0143 | 0.0184 |
| f | f | f | 0.1214 | 0.1173 |

The list of conditional probabilities in the Bayesian network is now also simplified into:

P(MotifsDetected) = 0.8643
P(CalculationOK | MotifsDetected) = 0.6694
P(CalculationOK | ~ MotifsDetected) = 0.0
P(Success | CalculationOK) = 1.0
P(Success | ~ CalculationOK) = 0.1356

The simpler Bayesian model is in fact a good approximation to the full Bayesian model. Table 1 gives a comparison between complete joint probability distribution computed with the full model and with the simplified one. Mean absolute difference between full model's joint probabilities and simplified model's joint probabilities is 0.0021. Mean absolute difference relative to the average probability value in this distribution (i.e. 0.125) is 1.6 %.

## 5 Discussion

The following observations are indicated by our experimental results:

- The results largely confirm de Groot's problem solving model. The first stage of problem solving is concerned with the detection of motifs, and the second stage is devoted to the calculation of concrete variations to verify whether motifs can actually be exploited. In all the 19 cases when relevant motifs were not detected, the calculation of variations was incorrect. It seems that it is practically impossible to calculate variations correctly without relevant motifs. When the motifs were detected, the calculation was at least adequate in about 67 % of the cases.
- In rare cases, about 10 %, a player did manage to solve the problem successfully even without detecting the motifs and with incorrect calculation. This can be explained by the fact that the number of generally reasonable moves in a chess position may be rather small, so even a random choice may succeed occasionally.

- Relevant motifs were correctly detected in 121 cases, that is 86 % of the total of 140 cases. In these 121 cases, the calculation was completely correct in 43 % of the cases, and calculation was at least adequate in 67 % of the 121 cases. This indicates the conclusion that for our players and type of chess problems, it was easier to detect the motifs than to perform the calculation.

This last point indicates that in our experiment, the dominant discriminating problem solving component between the players' success and failure was the calculation of variations and not detection of motifs (that is domain specific knowledge). Not much difference, in terms of players' problem solving performance, arises from the differences in players' pattern-based knowledge. This may be surprising because our players' chess strength (measured in Elo ratings) varied so much. The results show, possibly against the common intuition, that it is the calculation ability that decisively differentiates between players' chess problem solving success and failure.

To interpret this finding carefully, it should be noted that in our experiment, we used *tactical* chess problems. In contrast to long term strategic problems (also referred to as "positional play"), tactical problems are expected to require more calculation. However, as our results show, calculation alone without pattern knowledge (motifs) is far from sufficient.

Statistical evidence supporting this conclusion regarding pattern knowledge vs. search is also reflected in the correlations between variables. The correlation coefficient between the players' correctness of solutions and the players' success in detecting relevant motifs is 0.4368. On the other hand, the correlation between the correctness of solutions and at least adequacy of calculation is much higher: 0.8870. It should be admitted that the correctness of calculation was, for each player and position, evaluated somewhat subjectively by a chess expert as explained in Sect. 3.1. However, this does not change the overall conclusion that follows from the experimental data. Namely, that the ability to calculate variations is more discriminative among the players regarding their success than is the ability to detect relevant motifs.

The conclusion about calculation ability being more discriminative than pattern knowledge is relevant to a discussion in chess of different views regarding chess teaching and training even at top grandmaster level (Kotov 1971; Wenzhe 2001). The question is which component of chess skill is more important and deserves more attention in teaching and training: chess-specific pattern knowledge, or the ability to calculate variations. There is a common agreement that both of these components are necessary to play really well. However, one side of this discussion, including the Soviet school of chess (Kotov 1971), puts more emphasis on the deep understanding of chess, which includes chess-specific pattern knowledge. The other view is that deep understanding of chess is not sufficiently effective in practical game playing without very reliable support from calculation. Therefore some put relatively more emphasis on the ability to calculate concrete variations. The results of this paper might be interpreted as supportive of the latter approach in respect of this particular dilemma.

The differences between the two approaches are also reflected in chess training of very strong players. This is well illustrated by the following comments by the Russian grandmaster, a former world champion and successful chess coach Alexander Khalifman in the article “Le Quang Liem and the Soviet school of chess” (Chessintranslation.com 2011). Khalifman commented on the young Vietnamise player Le Quang Liem, the surprising winner of a very strong chess tournament in Moscow in 2011. Khalifman explains: “What a school is and what its presence or absence means is something that you can understand very well if you analyse with Asian chess players…. I worked a little bit with Le Quang Liem, and I will say honestly that sometimes my eyes popped out of my head. He is also a very talented boy… and he is trying very hard to grow. But at the moment all he does is calculate and calculate variations. He calculates very well, by the way. But a school is, in my opinion, what you would call a basis of positional principles, playing from general considerations…” This is how Khalifman described his view on the relative deficiency of elements of Soviet chess school in the young Asian player who was, in Khalifman’s opinion, overwhelmingly relying on calculation.

## 6 Conclusions

Our results indicate that the dominant discriminating problem solving component is the calculation of variations and not detection of motifs. This may come as a surprise in the view that traditional chess teaching puts so much emphasis on general chess knowledge which includes detection of chess motifs.

In our experiments, no significant difference in players’ problem solving performance arises from the differences in players’ pattern-based knowledge (ability to detect motifs), despite large Elo rating differences between the players.

Our results also clearly confirm one aspect of classical De Groot’s model of chess thinking. Namely, that satisfactory calculation is not possible without detection of motifs.

It should be noted that these findings are limited to the type of problems used in our experiments, that is tactical chess problems. One question for future work is whether the relative importance of calculation vs. chess pattern-knowledge also extends to non-tactical positions, particularly to long-term positional play. A first intuition on this might be that in such chess positions calculation of variations is relatively less important than in tactical positions. On the other hand, even in sharp positional play it is important to calculate how positional motifs can be realised by concrete sequences of moves.

Another topic of future work is to develop a program for automatically detecting players’ motifs directly from the players eye tracking data. This task seems to be rather demanding, but it could help to reduce some uncertainty in interpreting players’ retrospections regarding the detection of motifs.

## References


De Groot, A. D. (1946). *Het denken van den schaker*. Mij: Noord-Hollandsche Uitg.

De Groot, A. D. (1978). *Thought and choice in chess,* Vol. 4. Berlin: Walter de Gruyter.

Elo, A. E. (1978). *The rating of chess players, past and present*. New York: Arco Pub.

Glickman, M. E. (1999). Parameter estimation in large dynamic paired comparison experiments. *Journal of the Royal Statistical Society: Series C* (*Applied Statistics*), *48*(3), 377–394.

Hristova, D., Guid, M., & Bratko, I. (2014a). Toward modeling task difficulty: the case of chess. In *COGNITIVE 2014, the sixth international conference on advanced cognitive technologies and applications* (pp. 211–214). IARIA.

Hristova, D., Guid, M., & Bratko, I. (2014b). Assessing the difficulty of chess tactical problems. *International Journal on Advances in Intelligent Systems, 7*(3 & 4), 728–738.

Kotov, A. (1971). *Think like a grandmaster (english translation)*. London: B.T Batsford.

Newell, A. (1990). *Unified theories of cognition: The William James Lectures 1987*. Cambridge: Harvard University Press.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Upper Saddle River: Prentice-Hall.

Reingold, E., & Charness, N. (2005). Perception in chess: Evidence from eye movements. In G. Underwood (Ed.), *Cognitive processes in eye guidance*. Oxford: Oxford university press.

Tikhomirov, O. K., & Poznyanskaya, E. D. (1966). An investigation of visual search as a means of analyzing heuristics. *Soviet Psychology*, *5*, 2–15. (Translated from *Voprosy Psikhologii*, *2*, 39–53.).

Wenzhe, L. (2002). *Chinese school of chess: The unique approach, training methods and secrets (english translation)*. London: B.T. Batsford.


## Online Resources


Chessintranslation.com, 02.03.2011. http://www.chessintranslation.com/2011/03/le-quang-liem-and-the-soviet-school-of-chess/. Last access January 1, 2016.

Chess Tempo web site. www.chesstempo.com. Last access January 6, 2016.