Human chess skill

2

Neil Charness
Wilfrid Laurier University

Should a computer be more like a man?

At one point in *My Fair Lady*, Henry Higgins cries out in exasperation: "Why can't a woman be more like a man?" Higgins, as you may recall, was an expert linguist who had taken on the seemingly impossible task of transforming an uneducated Cockney flower girl into a very proper high society lady. The task was a close analogue to the classical Turing test.[1] The requirement was really that Higgins' protege successfully *simulate* the behaviour of a high society lady such that no observer would suspect otherwise.

No doubt many computer programmers have uttered similar cries of anguish in the process of getting their proteges, computer chess programs, to reach equally high standards of excellence in chess play. Should a computer be more like a man when it comes to playing chess? The argument for a more exact simulation of human chess players has been forcibly advanced by former World Chess Champion, M.M. Botvinnik [18]. He has suggested that since humans seem to be suited quite well to solve "inexact" tasks like choosing a move in a chess game, and computers are not, teach the computer to behave like the human.

On the other hand, a second school of thought is to consider that computers, because they operate differently than humans (at least on the "hardware" side), might perform better by adopting procedures that are maximally efficient for them (but not necessarily for humans). Thus one

---

[1] A. M. Turing [95] proposed a test in 1950 to decide if machines could think. The test involved having a human and a computer in separate rooms answer questions asked by an observer who could only communicate on a teletype. If the observer could not distinguish which was which, the computer was deemed capable of thought.

could imagine a successful chess playing program which would not play the same way a human does. Indeed, no current program does, and some of these programs play a reasonably good game of chess.

Perhaps the clinching argument in favour of the first line of thought is that if your goal is to develop a program which can beat the best human player, a guaranteed solution is to have the program simulate human playing methods almost exactly. This way you can capitalize on the computer's "inhuman" ability to avoid making mistakes in calculation.

The reason for this not so idle speculation about humans and computers lies in the failure of some predictions that have been advanced about computer chess. Not too many years ago, when the first generation of chess-playing programs appeared, some rather dire predictions were made. Before too long the best chess players in the world would be nonhuman, that is, computers.[2] The present generation of programs have not even achieved the more modest goal of playing master-level chess. If major advances do not occur between now and 1978, several professors of artificial intelligence are in danger of losing a somewhat weighty bet with David Levy, an international master chess player [67].

What has gone wrong? I think it is fair to say that the problem has little to do with "debugging" already capable chess programs (correcting programming errors). Rather, the primary reason for lack of progress has to do with our failure to appreciate the complexity of the task and our lack of understanding of how the human chess player performs the task. What characteristics do good human players possess that poorer players and programs do not? How does the human go about solving the problem of choosing a move?

The choice-of-move problem

Chess is a game that involves the alternation of moves by two armies: White and Black. The goal of the game is to checkmate the opposing king. This is accomplished when the enemy king is placed in check by one (or two) opposing pieces in such a position that escape is impossible because all flight squares are occupied by friendly pieces or controlled by opposing ones and when there is no possibility of capturing the checking piece(s) or blocking its (their) attack.

Few games are won by the strategy of trying to reach such a "mating" position directly. Instead, in high-level games, players strive for more modest goals like winning material or acquiring a positional advantage, knowing that it is possible to transform such "winning" positions to mating positions. Thus, in choosing a move from an initial position, the player attempts to transform it into one where he obtains a winning advantage, or one where he more closely approximates that goal.

---

[2] One such overly optimistic prediction was made by Simon [86] when he predicted in 1958 that a computer program would be world champion within 10 years.

Since chess is a finite game with a limited number of legal positions, it is possible, in theory, to generate all possible legal continuations. This can be done by generating all legal moves, the opponent's legal replies, the counter-replies, etc., until terminal positions are reached where either the White king is checkmated, the Black king is checkmated, or a drawn position has resulted. Then, by tracing his way back through the "tree" of possible moves, the player can find the best move for his side by using the so-called "minimax" procedure. He chooses the path which maximizes his outcome, given that both he and his opponent choose the move with the highest value for their respective sides at each level in the tree. (See Frey's chapter for a more detailed analysis.)

A quick piece of mathematics will convince the reader that the approach of exhaustive search through all legal positions is not a particularly efficient way to play chess. If from an initial position there are 10 legal moves for you, 10 legal replies for your opponent to each of your moves, 10 legal counter-replies to each of his replies, etc., then to look ahead only 6 half-moves (plies) you would need to evaluate $10^6 = 1$ million terminal positions. When you consider that in a typical master level game there are on average 42 moves (84 plies) in a game and an average of 38 legal moves per position,[3] the number of positions which could theoretically be explored is 38^{84}.

But as de Groot has pointed out [32], the number of master game positions that could arise is somewhat more constrained. He has estimated that the average number of good moves in a given position is 1.76. Also, the average number of moves in a master game is 42—i.e., 84 positions. Thus it seems that the number of positions to be explored in an "exhaustive" master search shrinks to 1.76^{84} or about 4.2×10^{20} positions. To illustrate how big that number is, consider that less than 10^{18} seconds have elapsed since the earth was formed some 4.6 billion years ago. Obviously, a chess player must search *selectively* among the exponentially exploding number of positions as he attempts to look ahead.

No doubt, many misconceptions about chess mastery were inspired by recognition of the difficulty of the search problem. One popular belief was that the highly skilled player calculated many more moves ahead than the amateur and hence could foresee combinations leading to winning positions which his opponent could not. Another view held that masters examined many hundred of continuations before choosing a move. The poor amateur, who could scarcely keep track of a handful of moves, would of course be no match for the formidable "thinking machine" called the chess master. Some of these myths were accepted by early investigators of human chess skill (see, for instance, Binet [15]). As we will see later, these misconceptions were demolished by de Groot and his coworkers in Holland. A simple refutation can be easily demonstrated by altering the depth of search for current chess programs. Adding a few plies does not

---

[3] See de Groot [31, p. 14, 19] for a derivation of these values.

produce tremendous changes in performance. What is gained by adding more look ahead capacity is rapidly lost in time spent evaluating the hundreds of thousands of new terminal positions. The only feasible solution to this exponential explosion is to use a highly selective search through the relevant branches. But how does one decide what is relevant and what is not?

The role of perception

"He saw everything!" is invariably the complaint of the chess player who loses a game. Other variants to this lament are: "I completely missed (seeing) his move" or "How could I overlook that move?" It is no accident that the operation "seeing" is an element in all those statements. In the final analysis, *perception* seems to be the key to skill in chess.

It is not usually the case that one player calculates so many variations that he generates the correct one where his opponent, who has searched less completely, does not. As is obvious from the previous analysis, both players are restricted to looking at a mere handful of possible positions. The difference between two players is usually that one looks at the promising moves, and the other spends his time going down blind alleys. This, in a nutshell, is what de Groot discovered in his research into the determinants of skill in chess in the early 1930's and 1940's.

At first de Groot tried to determine why some players were better than others by examining their thought processes. To do this, he showed chess players unfamiliar positions and told them to choose a move and think out loud while doing so. He recorded their verbal statements by hand and attempted to estimate their statistics of search. His subjects ranged from the world's best chess players—Grandmasters like Alekhine, Keres, Euwe—to club players who would be rated as Class A to Class C on the USCF rating scale.[4] Table 2.1, derived from de Groot, gives the results for position A (see Figure 2.1) by five Grandmasters and five experts.

The results are rather surprising and serve as a convincing refutation for the myths mentioned earlier. The only measure which clearly differentiates the Grandmasters from the experts is the one giving the estimated value of the chosen move. Four of the five Grandmasters chose the objectively correct move. None of the experts picked that move. Furthermore, as de Groot shows elsewhere [31, p. 128], all of the Grandmasters mentioned the correct move at some time in their analysis. Only 2 of the 5 experts ever mentioned the correct move in their analysis.

---

[4] Arpad Elo [36] has derived a method for rating players by comparing their performance against each other in tournaments. In the USCF (United States Chess Federation) system, the standard deviation is 200 rating points and the scale runs as follows: Class E: below 1200, Class D: 1200–1399, Class C: 1400–1599, Class B: 1600–1799, Class A: 1800–1999, Expert: 2000–2199, Master: 2200 and above. Titles such as international Grandmaster (approximately 2500) are awarded by FIDE, the world chess body.

TABLE 2.1 Choose a move task
Position A (After de Groot [31])

Protocol structural variables	Grandmasters (N = 5) Mean	(Standard error)	Experts (N = 5) Mean	(Standard error)
1. Time to decision (minutes)	9.6	(1.50)	12.8	(2.85)
2. Number of fresh starts	6.6	(1.75)	6.4	(1.57)
3. Number of different first moves	4.2	(0.58)	3.4	(0.68)
4. Maximal depth calculated (plies)	6.8	(0.91)	6.6	(1.36)
5. Total number of move considered	35.0	(10.66)	30.8	(8.14)
6. Moves/minute [(5)/(1)]	3.5	(0.59)	2.5	(0.53)
7. First moves/minute [(3)/(1)]	0.45	(0.05)	0.29	(0.04)
8. Value of chosen move (best move rated 9)	8.6	(0.40)	5.2	(0.20)

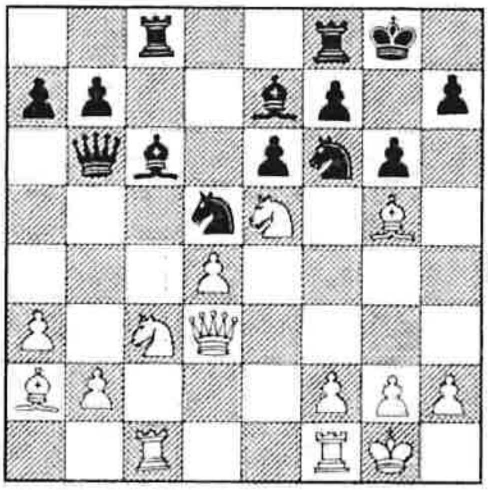

Figure 2.1 De Groot's position A with White to play. The best move is BxN/5!

It is also enlightening to observe that the average depth of search, for Grandmaster and expert alike, has shrunk from mythical levels such as 40 plies to a mere 7 plies. Also, the "hundreds" of continuations supposedly explored has dropped to an average of 35 moves in total (the range being 20–76 for Grandmasters).

Quite clearly, there was something about position A that attracted the Grandmasters' attention to the correct move, but not the attention of the experts. An argument can be advanced that the two groups *perceived* the position quite differently.

The role of perception became much better defined when de Groot modified a task used in an early Russian investigation [33]. He again permitted his subjects to look at a chess position taken from an unfamiliar game but limited the exposure to a few seconds. Then, after the players had reflected upon and organized what they had seen, they were encouraged to recall the position. Stronger players did so by calling out the positions of the pieces verbally. Weaker players were permitted to place pieces on an empty board.

The results of this experiment were striking. Grandmasters and masters recalled almost all the pieces correctly—approximately 93 percent of 22 pieces. Experts recalled 72 percent correctly, and class players recalled about 51 percent correctly.[5] William Chase and Herbert Simon [22, 23], at Carnegie-Mellon University, replicated and extended this finding by adding a novice to their group of subjects and carrying out a valuable control procedure. Their master scored 81 percent, the A player 49 percent, and the novice 33 percent with a 5 second exposure to "quiet" middle game positions taken from master games. But when randomized (scrambled) versions of chess positions were shown for 5 seconds, everybody, including the master, recalled only 3 or 4 pieces correctly—fewer pieces than were recalled by the novice with structured positions.

Chase and Simon were thus able to conclude that the superior performance of the master was not due to extraordinary visual memory capacity, but rather to a chess-specific capacity. This conclusion parallels earlier ones by Djakow, Petrovsky, and Rudik [33] and de Groot [31]. These studies in perception suggest it is not the thought process but the perceptual process that seems to differentiate chess players according to skill level.

Because recall ability seems to be the most sensitive measure related to chess playing ability, it becomes important to discover how players perform the recall task. Simon and Gilmartin [85] have produced a model of the task that is embodied in a computer simulation called MAPP.

In this model, Simon and Gilmartin make a series of assumptions about the information processing capacities of humans. They assume, on the basis of much psychological research, that people possess two memory systems: short-term memory (STM) and long-term memory (LTM).[6]

Human memory

Short-term memory apparently has a limited capacity. This memory system represents what you are currently aware of, and it is therefore sometimes referred to as working memory or immediate memory. The limited capacity aspect of STM is best illustrated by the difficulty people have remembering a new, unfamiliar telephone number which they have just looked up. Unless they rehearse the 7 digits, they are likely to forget them before they can dial the number. Furthermore if people are given 2 such numbers in quick succession they have virtually no chance of recalling both of them accurately.

This observation does not prove that the capacity of STM is 7 digits. If the two phone numbers are familiar ones (e.g., home and office), most people have no difficulty at all remembering them, but they do run into

---

[5] De Groot used a rather complicated scoring procedure whereby he awarded points for pieces placed correctly, and as well, for remembering spatial relations between pieces and material balance. He subtracted points for misplacing, adding or omitting pieces, as well as for interchanging pieces, shifting them over one file, and being uncertain. (See [31, p. 223].)

[6] There is also good evidence for the existence of a very short-term memory. This memory holds information sent from the sensory system for about ¼ of a second.

trouble if they are given 7 familiar phone numbers. From various experiments, George Miller [72] concluded that the capacity of STM is 7 *chunks*, plus or minus two. A chunk is a psychological unit—a familiar unit to the person. It is a shorthand code which can later be decoded to recover all the information it represents. It is essentially a label or symbol which designates or points to specific information in long-term memory.

Short-term memory is sometimes conceptualized as a storage bin with seven locations or slots. The organization of STM is temporal. Items can be entered serially as each is attended to. Once all seven slots have been filled, however, something must be lost if new information is to be processed. Usually the oldest item is replaced by the newest. Fortunately we have the ability to rehearse the contents of STM and the capacity to recode information and store it in a more permanent memory: LTM. Storing information in LTM takes time. Simon [83] has estimated that it takes about 5–10 seconds per chunk. Because of its small capacity and relatively long transfer time, STM is the major bottleneck in our ability to process information.

Long-term memory is the permanent storehouse of all that we know. It is virtually unlimited in capacity—barring certain neurological disorders, we can continue learning until we die and never exhaust our storage capacity. Items such as your phone number, your name, and the alphabet are stored in LTM. Information that is used frequently or which is intentionally practised or recoded resides permanently in LTM. Information in LTM is used to interpret and restructure information transmitted from the senses to STM.

A fascinating property of LTM is its organization. Information in LTM seems to be highly interconnected. This property contrasts sharply with most computer memory systems, where information is said to be location addressable. That is, one piece of information is linked to another by a location tag: e.g., "go from this cell (300) and get the contents of cell 385." Barring the existence of other pointers to cell 385, the only other way to retrieve its contents is via an exhaustive search of all memory cells —a highly inefficient way to retrieve information. The only practical way to get that information is to start from cell 300. What happens if the entry cue is changed slightly such that it no longer activates cell 300? The information in cell 385 cannot be retrieved.

Human LTM appears to be content addressable. That is, functionally similar items seem to be filed in the same location.[7] Items may be grouped together on the basis of semantic similarity (meaning), phonetic similarity (sound) or other categorizations. An example of successful retrieval based on semantic similarity occurs for most English speakers when they are

---

[7] Location is not meant to refer to an exact place in the brain. Memory capacity is apparently diffused throughout the brain, with some exceptions: e.g., language information which, for most people, is present primarily on the left side of the brain. Location thus refers to functional location of information—which may mean a series of neural pathways not occupying the same physical location.

asked for synonyms for "speedy." In short order people often report back words like "quick," "rapid," "swift," "fast." Your ability to retrieve information via phonetic cues can be demonstrated when you are asked to name words which rhyme with "bat." Words and concepts are not cross-listed under every possible category, however. For instance, most people have considerable difficulty in responding quickly to the demand "name words whose fourth letter is 'a.' "

When the memory system is confronted with a cue which is not directly associated with the desired information, it can often search the general area where the information is stored. This type of search is really a form of problem solving. Thus there are many possible paths to get from one piece of information to another, even when the items were never activated together before: e.g., name a type of dog that rhymes with "folly."[8]

Despite the highly connected structure of LTM, there are many occasions when information may not be accessible. There is a distinction between inaccessibility and absence. This distinction often underlies the difference in sensitivity between recall and recognition. "Who was Lyndon Johnson's Vice President?" Although many readers may not know the answer immediately, they may "know" that they know the answer and this knowledge may induce them to do a prolonged memory search. If the question were rephrased as a recognition task—Was Nelson Rockefeller, Richard Nixon, Hubert Humphrey, or Henry Kissinger, Lyndon Johnson's Vice President?—many more readers will select the correct answer, thereby demonstrating that the information was present in memory.

On the other hand, people often know that they do not know an answer and refuse to do much searching at all: "What is the first name of this writer's wife?" The reader, on analyzing the question might have gone to the location in memory where this information, if it existed, would be found and discovered that there was nothing present under that category.

Obviously, the more information there is stored in LTM, the more things can be recognized and labeled, provided that the information is stored in an organized fashion. Simon and Gilmartin postulated in their model that better chess players have more chess patterns (which they can thus recognize) stored in LTM than poorer players. In the short time available for looking at a chess position in the de Groot task (5 seconds), they assume that everybody is restricted to storing 7 chunks in STM.[9]

---

[8] One way this problem might be solved is by a *generate and test* method. Generate a new letter for the "f" and then test the resulting sound pattern to see if it forms a word in the dog category. See Newell and Simon [75] for an excellent discussion of methods of problem solving.

[9] Given the previous estimate of 5–10 seconds per chunk fixation time, this assumption is quite reasonable. The main problem with this parameter is that it was estimated from experiments in learning verbal materials. Charness [20] has obtained evidence suggesting that virtually all the information extracted during a 5-second exposure to a chess position is stored in LTM. This issue has not been satisfactorily resolved yet.

The key to the mystery of performance in the 5-second task is provided when you consider that the master's average chunk size may be 3–4 or more pieces, e.g., the castled king position in Figure 2.1, whereas the novice's chunk is a single piece on a square. Thus 7 × 3 = 21 pieces for the master, but 7 × 1 = 7 pieces for the novice. The idea that everyone has the same-sized STM is given further support by the condition of presenting scrambled chess positions. In that situation there are no longer any recognizable patterns larger than single pieces for the master, and he sinks to the level of the novice.

An analogy to the chess board situation for master versus amateur is the case of a child learning to read versus an adult who is already a skilled reader. When the child looks at this page he "sees" a page filled with letters which he must slowly and effortfully recombine and read as words. The adult, however, quickly and effortlessly "sees" the page as a series of words and possibly phrases (which may then have to be effortfully recombined into sentences compatible with his current knowledge about human chess skill). Both the adult and the child look at the same page but they produce very different encodings or descriptions of it, based on the size of the pattern they can use for effortless recognition.

The specific model developed by Simon and Gilmartin is considerably more complex than is outlined here—it describes how patterns are originally learned and later recognized, which pieces will be attended to, how the patterns are reproduced—but the reader can easily perceive its explanatory power.

The simulation does a very good job of imitating strong and weak players by simply altering the number of patterns which are stored in LTM. Simon and Gilmartin estimated the size of the vocabulary of patterns that a master would theoretically need to perform the recall task as well as he does. They arrived at an estimate of about 50,000 patterns—which is roughly the same size as the vocabulary of recognizable words for an adult speaker of English.

Does this research mean that all one needs to do to become a Grandmaster is to sit down and memorize 50,000 chess patterns? I can hazard a guess that if you attempted to do this you would undoubtedly perform as well as the Grandmaster on the recall task, but your chess play would hardly improve. How then is it possible to link up this vast knowledge of patterns with chess skill?

Chase and Simon have suggested that the correlation between perceptual skill and chess skill can be accounted for by assuming that many chess patterns are directly associated with appropriate plausible moves. That is, when the master looks at a chess position his recognition of familiar configurations of pieces triggers certain "productions" into action. A production is a behavioral unit which has two components: a condition side and an action side. It can be modeled by a statement of the form: if condition *X* exists, do action *Y*. A chess production might be: if pattern *X*, consider move (plan) *Y*; or, more concretely: if there is an open file, consider moving a rook to it. This can be illustrated by way of a more complex example. Most skilled chess players will recognize the smothered mate position illustrated in Figure 2.2. White, if on move, can mate in 4 moves (or less) via (1) N–B7 ch., K–N1; (2) N–R6 dbl. ch., K–R1 (if K–B1, Q–B7 mate); (3) Q–N8 ch., RxQ; (4) N–B7 mate. Changing nonessential elements of the position (moving the black QRP to QR3) does not change the mate. Other changes, however, make a big difference: e.g., moving the black KRP to R3 or interchanging the black queen and rook.

Humans are probably sensitive to the critical features of such a position. They do not store a copy of each possible smothered mate position or each back row mate position, etc. They probably abstract more general descriptions. In the case of this type of smothered mate the features are probably propositions like: queen on open QR2–KN8 diagonal; knight capable of reaching KB7 in one move; opponent's king on KR1 hemmed in at KN2 and KR2. Probably the latter feature together with one of the former is sufficient to trigger the plan: try to reach a smothered mate position. One can come up with many other examples of such typical tactical plans which are part of every skilled player's repertoire.

Now there is a potential explanation for why, in position A (Figure 2.1), the Grandmasters all considered the correct move, but only a few experts did. There were probably certain features in the position which quite automatically, *when recognized,* elicited the appropriate move. Only those players who recognized the features and possessed the appropriate productions would generate the correct move for subsequent evaluation.

Parenthetically, it also becomes reasonable to speculate about questions like: why does a "highly intelligent" individual when playing chess, miss obvious moves? Moves are only "obvious" when the patterns they spring from are recognized.[10] One can also ask why masters do so well in

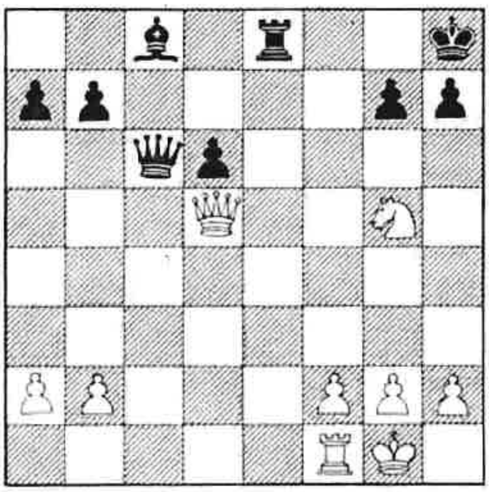

Figure 2.2 A smothered mate position with White to move.

---

[10] See Botvinnik's text [18] concerning Norbert Wiener, p. 61, for another explanation.

simultaneous exhibitions,[11] where they have only a few seconds to choose a move. If masters automatically generate appropriate plausible moves, these moves will usually be good enough to beat all but the best players at the exhibition. A similar explanation will probably also suffice to explain a master's superiority at speed chess—where the entire game is played in less than 10 minutes. Indeed, if the perceptual process is so important for skilled play, it becomes possible to understand the idiosyncracies of a Bobby Fischer, who insists that the lighting in the tournament hall, the size of the board and the size of all the pieces, all be optimized.

The first few seconds

Following the analysis of de Groot, and Chase and Simon, much of what is critical to choosing a move in chess appears to occur in the first few seconds of exposure to a new (or changed) position. In those first few seconds the skilled player has constructed his internal representation of the position. This phase is usually accompanied by a series of eye movements which fixate on parts of the board or diagram. As Tikhomirov and Posnyanskaya [92] and others have shown, these fixations are made on the functionally important or salient pieces. That is, in the first few seconds when an expert is told to choose a move, he first explores the relations which bind the pieces into functional units, i.e., chunks. Only later does the pattern of eye movements trace out the paths of pieces when they are moved in the "mind's eye." What are the relations which bind pieces into a chunk? Chase and Simon examined the pattern of pauses in a simple perceptual task. The task required that chess players reconstruct a position seen on one board onto a neighbouring board. Players had to glance back and forth between the two boards (presumably to transfer pieces chunk by chunk), and their head movements and reconstruction were recorded on videotape. It was found that chess players paused significantly longer when placing a new piece if the piece had fewer relational links with the previously placed one than if it had many such links. These relations were of two types: chess functional like attack and defence and visual-spatial like proximity, color, and type of piece. If two successively placed pieces had a large number of these relations in common, the latency between their placement was quite short. The function relating interpiece latency and number of shared relations is shown in Figure 2.3.

Thus these researchers were able to use the criterion of a latency greater than 2 seconds to indicate a chunk boundary. Apparently, even during the first eye movements, the skilled chess player is sensitive to important functional relations among pieces.

This phase of perceptual organization was first postulated by de Groot

---

[11] In a simultaneous exhibition the master plays many opponents at once, usually 20 or more, moving steadily from board to board, playing a move at each board.

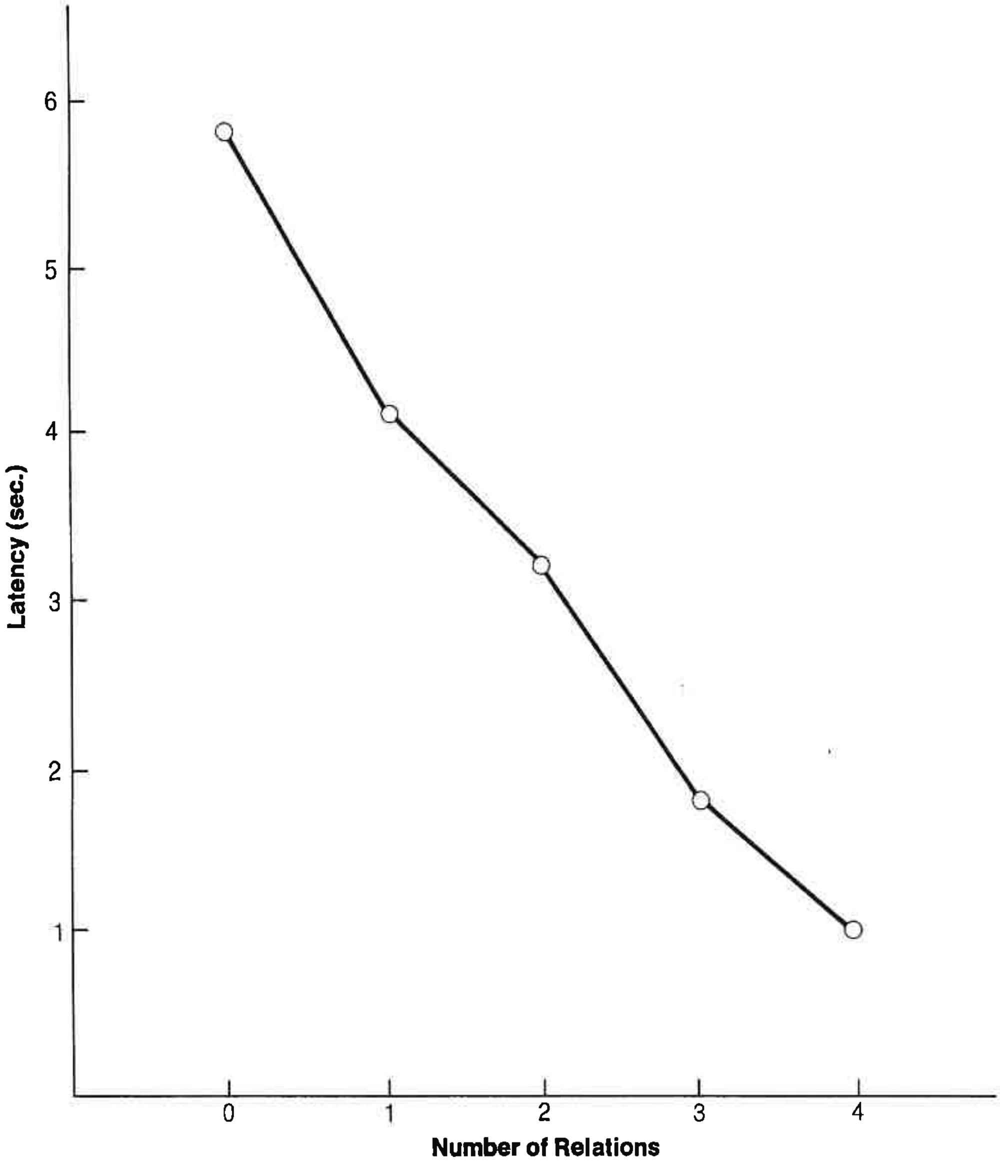

Figure 2.3 Research data on interpiece placement latency in a task where the subject is asked to set up a chess position while viewing the position. (Adapted from Chase and Simon [23], with permission.)

when he noticed that players in his choose-a-move task were initially silent. De Groot notes that it is during this time that the player obtains the "sense" of the position. He recognizes patterns, discovers the state of material balance,[12] and starts generating plausible moves or plans. Following this phase, the thought processes look quite similar for master and less-skilled player, although the paths they explore are quite different.

---

[12] Material balance refers to how many pieces are present in each army. A position is considered balanced when the same number of pieces of each type are present, or if there is equivalent material present.

Search through the tree of moves

Much of what we know about the analysis procedure which players use in evaluating a move comes from the work of de Groot and that of Newell and Simon [75] and Wagner and Scurrah [98]. Briefly, they found that players explored their candidate moves one branch at a time by a method called progressive deepening. (These authors all used the same procedure: having their subjects think out loud and recording their statements.) Players choose an initial move for consideration and then generate the consequences to some depth, usually no more than 6 or 7 plies. They may then reexplore the same base move by returning to the initial position and then considering new countermoves further down the line. Opinions differ on the rules for generation of a search episode (or line of moves), but they all seem to hinge on the evaluation which is extracted from the terminal position.

One aspect of progressive deepening which seems quite nonintuitive is that the same move may be taken up again later in search. This resumption of an abandoned line is not necessarly the result of a player's having forgotten his analysis. Rather, it indicates that the player may have changed his conceptualization of what is required in the position. Thus a move which is rejected because it failed to lead to the win of material may again be taken up in the context of trying to gain a positional advantage. Much of the human's search through the tree of possible moves is goal directed and the goals may well change as new information is integrated into the knowledge base about the initial position.

Virtually every line of moves mentioned by a player in these experiments was terminated by an evaluative statement. That is, the players noted an advantage or disadvantage for the line of moves. Statements were made like: "not sufficient," "that seems to win a piece," "(the move) will probably win" (see de Groot [31, p. 409]). Of course, that is the goal of dynamic analysis—to reach an evaluation of a move by exploring and evaluating some of its outcomes. One hopes that the "some" in the previous sentence encompasses all of the relevant continuations.

Terminal positions in search are evaluated statically. Some feature or property of the position is usually cited in the evaluative comment, usually just one feature. In this respect, virtually all chess playing programs differ radically from the human. Machines usually rely on a polynomial weighting function to evaluate a position. That involves counting up the pros and cons of a large number of features such as material balance, mobility of pieces, safety of the king, and control of the center. It may be the case that the human conceals the calculations which he goes through to reach an evaluation, and merely mentions the feature with the greatest weight. At present, though, there is no evidence to suggest that the human does go through such an elaborate procedure. One possible reason why only one feature is mentioned is that the line of moves under examination has probably been generated in response to a particular board goal. The evaluation procedure is then reduced to determining if that goal has been met.

As was mentioned earlier, the moves that a player examines are not a random subset of all legal moves. To cut down on searching irrelevant moves, people (and some programs) search heuristically. That is, they use information about chess to help them generate and evaluate a restricted number of moves. Winning a knight for a pawn, all other things being equal, is known to be good. Similarly, any move which will win the "exchange" is evaluated positively. Other heuristics which help chess players focus their search are rules of thumb like: in the opening, move your king to safety by castling; control the central squares of the board; develop your pieces to places where they have high mobility. These principles are typically espoused in chess books.

Such heuristics are valuable because they provide general rules which can be filled by specific types of moves. On the other hand, they are too general to act as hard and fast rules which, when followed, enable a player to play well. That is, they do not act as *algorithms*[13] for playing chess.

Another type of heuristic process which is often found in human play is to make use of information generated in searching a specific move. A particular move found somewhere along a line of investigation may itself become a candidate as a base move. Or certain propositions about relations between pieces may be used to generate new plans or moves. For instance, a player may notice that in following one continuation that one of his opponent's pieces becomes tied down to the defence of another piece. This may elicit the idea of attacking the tied down piece in order to win the other piece—a theme called attacking an "overloaded" piece.

Visualizing positions

All this analysis is carried out in the player's head. How do players keep track of the new positions generated in analysis? Keeping track involves memory and some imagery and is susceptible to spatial errors. Occasionally a blunder can be traced directly to the "imagining" operation. A player will fail to "notice" that moving a piece to a certain square blocks another piece from giving check at a later critical moment. (For an amusing account of just such an hallucination see *Chess Treasury of the Air* [94, p. 273].)

The role of visual imagery in chess has long been of interest to psychologists. It was initially investigated to explain the ability of chess masters to play "blindfold" chess. In this situation the master plays a game entirely in his head—he hears the moves of his opponent (in chess

---

[13] An algorithm is a specific set of rules (i.e., a recipe) which, if followed exactly, will guarantee a successful result. The algorithm for long division is a good example of such a procedure. A particular long-division problem you face may never have been solved before, but you have no trouble doing it.

notation) and then replies with his own moves without ever seeing the board.

At first it was thought that the blindfold player had an image of the board much like that of a photograph or picture—namely that the shape of the pieces, the color of the squares, were all present in vivid detail. Binet in 1893 refuted this view after interviewing many masters. He concluded that the representation used by blindfold players was quite abstract. For instance, a player would know that there was a knight in a certain relation to other pieces in a particular position, but he would not have an image of a particular carved knight. He might know the knight was white, but he wouldn't *see* a certain shade of white.

The reader can easily convince himself that one does not form a detailed image of an entire position. Imagine an empty chessboard. "Place" a bishop on QR3. Name the farthest square the bishop can reach on the long diagonal. Now put a real bishop on a real board and do the same. The time taken to answer KB8 should be markedly different in the two cases. (Parenthetically, how many of you chose a white bishop? Or did the bishop in the imaging case have a color? Such *generative memory* is quite abstract.) Similar conclusions about the nature of the representation of the board were reached by de Groot, and by Reuben Fine in his article on blindfold chess [43].

Apparently the ability to imagine pieces on squares in a chess position is also correlated with chess skill. I conducted an experiment [20] where players were given chess positions verbally, piece by piece, via chess notation. After hearing a position dictated in this fashion at a fairly brisk pace, the players attempted to reproduce the position. Two groups of players were tested: Class A (mean rating = 1870) and Class C (mean rating = 1458). The order in which pieces were read made a large difference in recall, but beyond that, the A players recalled 14 percent more pieces per position, on average. That all players fared poorly when the pieces were "scanned" randomly indicates that they did not perform the task by "imagining" pieces on squares. Rather, they attempted to remember the pieces by remembering the relations which bound them together into chunks. The most favourable scan was one which consecutively named pieces with many relational links.

Evaluation

As was mentioned earlier, the reason for extensive dynamic analysis (search) by players was to evaluate the consequences of a candidate move. Static evaluation of the position after a move is almost never sufficient to judge it accurately. In essence the whole problem of choosing a move in chess comes down to choosing a move that can be correctly evaluated as being superior to others. Or, in a nutshell, the player must decide: does one move leave him with a better position than all others? The decision is extremely simple in the case of giving check and recognizing that it is mate. In most positions, however, the evaluation process is complex.

The point of searching deeply (many moves into a continuation) is to find a position which can be evaluated confidently. The stop rule for search is apparently just that—reaching a position which is recognized as static (or quiescent) and which can be evaluated as good or bad. The more skilled player—the one with more patterns in memory—should, in theory, be able to recognize a terminal position (dead position) sooner than a less skilled player. We *should* be surprised by de Groot's finding that Grandmasters and experts search as deeply into a position. If the above analysis is correct we might expect the Grandmaster to search *less* deeply before reaching a position which can be statically evaluated. Perhaps the apocryphal anecdote in this regard is the one attributed to Reti. When asked to tell how many moves ahead he typically looked when calculating a combination, he is reported as having said "as a rule not a single one" (see [94, p. 39]).

Again the realization occurs that the number of patterns stored in memory may be instrumental in understanding evaluation. It is probable that the chess patterns which are stored in memory are associated not only with information about plausible moves, but also with evaluative information. For instance, most skilled chess players rapidly evaluate the diagram in Figure 2.4 as a win for White. They undoubtedly recognize that in this type of position, with white a pawn ahead, white wins. But if you add either a pair of knights (one for each side), a pair of bishops, a pair of rooks, or a pair of queens to the position with the sole restriction being that no exchanges are forced, most skilled chess players will rate the position as drawn. In this instance, they recognize that in this type of position, with one side ahead a pawn but all pawns on the same side of the board, there is no forced win. Thus, for the skilled player, there is no problem in deciding to transpose to the former or the latter position when contemplating a series of piece exchanges in an earlier part of the game.

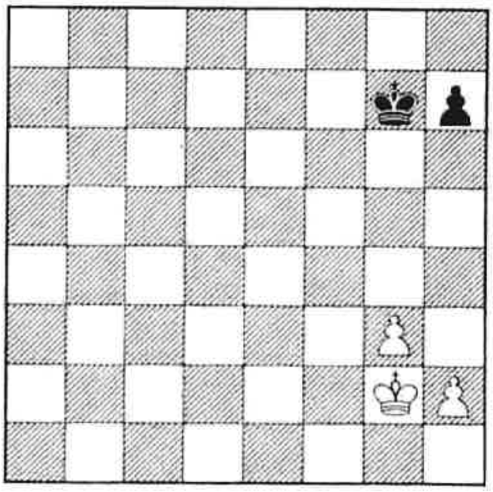

Figure 2.4 A position which is a clear win for White. If rooks, knights, bishops, or queens are present, however, and no exchanges are forced, the position would be a draw.

The player merely needs to generate these two terminal positions and recognize that one is a win, the other a draw. There is no need to calculate (tree search) any farther.

The more accurately a player can classify positions as positive, negative, or neutral, the more likely he is going to defeat an opponent who has less accurate evaluations—particularly if they both search the same continuations. (How many chess playing readers have been faced with the agonizing decision either to "close" a position or exchange pieces and open it up and have been unable to choose because they could not evaluate the resulting positions accurately? After the game, your opponent graciously points out that you should "obviously" have chosen the other alternative.)

It is possible, following this line of reasoning, to speculate on how masters make "intuitive" moves or sacrifices. Perhaps what is meant by an intuitive move is that the player has *recognized* that the resulting position is of a type where an attack almost always succeeds. He doesn't need to calculate out all the details. He knows that when he does search later for a combination he will find it.

Motivation

One elusive aspect of chess skill that psychologists have not come to grips with particularly well is the "will to win." Some players are characterized by their fierce desire to win at all costs. Others are content to seek draws and take few risks. How important is desire? What makes one player an attacker and another a maneuverer? If one player constantly looks for an opportunity to sacrifice material, he is of course more likely to head for positions in which this happens. Another who is more comfortable fighting in closed positions is more likely to play for those sort of positions.

One would have to predict, however, that any Grandmaster should play the same move in the vast majority of situations. This is of course implied by the fact that there is rarely more than one good move in a position, as de Groot has shown. Also, what we mean by chess mastery is that the player has to a large extent mastered the problem of choosing the best move.

There is some interesting data from Tikhomirov and Vinogradov [93] which suggests that if a player is not permitted to get involved in the game emotionally his problem solving ability decreases. Using galvanic skin response (GSR) as a measure of emotionality, these investigators discovered that as players neared a solution to a chess problem or discovered a critical line they showed an emotional response—a swing in galvanic skin response. Tikhomirov and Vinogradov then tried an interesting experiment. They asked the players to attempt to eliminate any emotion or arousal (control the GSR device) during problem solving. If subjects were successful in following these instructions, they solved problems of moderate difficulty but were quite unsuccessful with the more difficult ones. The inference, though speculative,[14] to be derived from this work is that unless a player is fully involved in a game, emotionally as well as cognitively, he may not perform at his best.

Thus there is some evidence to favour an oft-expressed opinion. Namely, that a player who has just arrived at a tournament from a previous one is unable to do his best because he is "all played out" and "emotionally exhausted." (A more sanguine explanation is that his opponents have looked at his recent games and prepared opening novelties against them.)

The road to mastery for man and machine

The process of choosing a move seems to involve perception as a primary component, and in particular, the recognition of thousands of stored patterns. It may seem surprising, but chess is not so much a game of deep thinkers as one of skilled perceivers. Perception apparently holds the key to understanding why the master barely glances at a position before seizing on the correct move, where the novice spends hours staring at a position and never generates the winning move.

Perception operates in two main areas: generating plausible moves and statically evaluating terminal positions. Not only will a plausible move be perceived more quickly by the master, but his long experience with many patterns on the board (50,000 or more) translates into the advantage of being able to evaluate the resulting position more readily.

The process of searching through the tree of possibilities, i.e., dynamic search, is surprisingly similar for both master and class player. They both search to a similar depth; they both examine about as many moves. The reason they end up with such different moves is that one goes down blind alleys whereas the other examines the critical variations.

How does the master build up this repertoire of patterns and chess-specific knowledge? Simon and Chase [84] suggest that it is a matter of practice. You cannot become a concert pianist by practicing one hour a day. Similarly you do not become a master chess player by thinking about chess one hour a day. Simon and Chase observed that it apparently takes a player nearly a decade of intense preoccupation with playing chess to become a Grandmaster—even child prodigies like Capablanca, Reshevsky, Fischer, and others were not exceptions.

Practice does not mean staring at and memorizing 50,000 patterns. It means learning to recognize types of positions and the plans or playing methods which go with them. The reason why the information necessary to become a master does not appear readily in chess books is that it is

---

[14] There is still much disagreement among psychologists over the interpretation of a change in GSR. In this case a change may be indicative more of concentration than emotion. It may reflect the degree of cognitive strain involved in searching many plies deep. See for instance Leedy and Dubeck [65].

primarily nonverbal; many patterns are difficult to describe in common language terms. Apparently, advice should take the form of: when you see this (pattern description), consider trying this plan. (Some current chess books are attempting to use this approach.) Also, the heuristics which appear in chess books are just not precise enough to cover the many concrete situations which occur over the board. Only the slow process of storing and classifying thousands of patterns seems to work.

As de Groot put it, for the master player "there is nothing new under the sun" ([31, p. 305]). Each position summons up typical playing methods or plans. Perception leads to playing method, which leads to search for the effects of specific moves.

This analysis may offer a rather discouraging picture to chess player and programmer alike. Unless somehow we can systematize and classify thousands of patterns and their appropriate playing methods, there is going to be no shortcut to mastery of chess. Kotov [60] has outlined what he feels are the fundamentals for becoming a Grandmaster. They can be summarized in five points:

1. Know opening theory
2. Know the principles behind typical middle games positions (methods and moves)
3. Be able to assess a position accurately
4. Be able to hit the right plan demanded by a given position
5. Calculate variations accurately and quickly

Almost all of these principles can be subsumed under three general principles: recognize patterns, know their value, and know their corresponding playing methods.

Will the computer reach and finally surpass the best of its human tutors? That day may be some time away unless some of the built-in advantages of humans are incorporated into chess programs. Humans have two major advantages over computers. They have both a vast knowledge base about chess and a set of procedures for efficient manipulation of that base. Master players possess an enormous amount of highly organized chess knowledge. They can recognize thousands of patterns and can evaluate them and generate plans for them. The latter point is an important one: human players are goal directed when playing chess. They know hundreds of heuristics that help structure their search. Computers also use heuristics but often the heuristics are not sensitive enough to the context of the board position.

The second advantage that humans have over most computer programs is the ability to modify their knowledge base. One aspect of this modification process is learning. Humans are learners *par excellence*. They continually add new information to their knowledge base. Chess players learn both specifics: e.g., "I'll never play that move in the Ruy Lopez opening again" and general principles: e.g., "To win this ending I must retain pawns on both sides of the board."[15]

Another type of modification that humans perform quite nicely (as do some complex programs: e.g., Winograd's language comprehension program [100]) is inference—inductive and deductive reasoning. This process is a form of problem solving which seems quite effortless for adults. Suppose you hear the following: "Eliot is a better chess player than Neil. Neil is a better chess player than Peter." Someone now asks you: "Is Eliot a better chess player than Peter?" Most adults rapidly answer "yes." Yet nowhere in either of those two sentences was there any stated relation between Eliot and Peter. By using your knowledge of relationships (that *better than* is a transitive relation for chess skill) you correctly inferred (deduced) that there was a link between Eliot and Peter. Inference is invaluable for chess play. It is this very process which underlies the "method of analogies" discussed in Frey's chapter.

Unfortunately, neither of these two capabilities is present to any sophisticated extent in present computer-chess programs. Thus the key to master level chess by computers would seem to be to provide them with a dynamic knowledge base—a large set of patterns, evaluations, and playing methods (board goals) which can be accessed, sometimes quite indirectly, and modified. Researchers have begun to make considerable headway toward understanding human chess play. If the principles of play uncovered by this research effort are implemented successfully in computer programs, the phenomenon of master level play by the computer may soon appear. Programmers ought to ask themselves: Why can't a computer be more like a man?

---

[15] That computer chess programs are not capable of such sophisticated self-modification need not dictate against the attainment of master level play. Programs function as models of performance. They can be considered to represent the knowledge state of a chess player at one instant in time. Programmers can step in to modify a program, thus acting as the learning component for the program. Admittedly this feedback loop is slower than the human's, but it is not in principle a limiting factor in reaching master level performance—eventually. Examine the other side of the coin: just because human chess players are capable of learning does not guarantee that they all become Masters.