Herbert A. Simon
William G. Chase

#383

Skill in Chess

Experiments with chess-playing tasks and computer simulation of skilled performance throw light on some human perceptual and memory processes

As genetics needs its model organisms, its *Drosophila* and *Neurospora*, so psychology needs standard task environments around which knowledge and understanding can cumulate. Chess has proved to be an excellent model environment for this purpose. About a decade ago in the pages of this journal, one of us, with Allen Newell, described the progress that had been made up to that time in using information-processing models and the techniques of computer simulation to explain human problem-solving processes (*1*). A part of our article was devoted to a theory of the processes that expert chess players use in discovering checkmating combinations (*2*), a theory that was subsequently developed further, embodied in a running computer program, MATER, and subjected to additional empirical testing (*3*).

Herbert A. Simon took his bachelor's and doctor's degrees at the University of Chicago, the latter in 1943. He has served on the faculties of the University of California, Berkeley, and Illinois Institute of Technology, and, since 1949, on the faculty of Carnegie-Mellon University, where he is Richard King Mellon Professor of Computer Science and Psychology. Beginning with an interest in decision-making in organizations, Professor Simon has been led during the past fifteen years into research on human performance in complex tasks, using the computer to simulate cognitive processes. With Allen Newell, he is coauthor of a recent book, Human Problem Solving.
William G. Chase is an Associate Professor of Psychology at Carnegie-Mellon University, where he has served since receiving his Ph.D. from the University of Wisconsin in 1969. His research has concentrated on the elementary information processes underlying cognition. He has edited a recent book, Visual Information Processing.
This research was supported by Public Health Service Research Grant MH-07722, from the National Institute of Mental Health. *Address of both authors: Department of Psychology, Carnegie-Mellon University, Pittsburgh, PA 15213.*

The MATER theory is an application to the chess environment of a more general theory of problem solving that employs heuristic search as its core element (*4*). The MATER theory postulates that problem solving in the chess environment, as in other well-structured task environments, involves a highly selective heuristic search through a vast maze of possibilities. Normally, when a chess player is trying to select his next move, he is faced with an exponential explosion of alternatives. For example, suppose he considers only ten moves for the current position; each of these moves in turn breeds ten new moves, and so on. Searching to a depth of six plies (three moves by White and three by Black) will already have generated a search space with a million paths. Hence, if *every* legal move is considered (as would be the case in an exhaustive search), an enormous search space would be generated. Such a search is beyond the capacity of the human player, as well as present-day computers. Humans seldom search more than a hundred paths in choosing a move or finding a checkmate, and they seldom consider more than two or three possible moves per position.

The MATER theory postulates that humans don't consider moves at random. Rather, they use information from a position and apply some general rules (heuristics) to select a small subset of the legal moves for further consideration. For example, one powerful heuristic that MATER uses in finding checkmates is to examine first those moves that permit the opponent the fewest replies. A comparison of the MATER program with thinking-aloud protocols from human chess players confirms the importance of heuristic search as a basic underlying process.

While the MATER theory was successful in accounting for much of what was known about chess thinking in mating situations, some important empirical phenomena—some of them known when the theory was formulated, some of them discovered subsequently—eluded the theory's grasp. In this paper, after describing the phenomena, we should like to tell the story of a ten-year effort to account for the recalcitrant facts.

An important by-product of this effort has been to bring about a convergence of the theory of problem solving with theories that have been developed to explain quite different phenomena, which psychologists label "perception," "rote learning," and "memory." In the past, both theorizing and experimentation relating to these different kinds of tasks—problem solving, perceiving, learning by rote, and remembering—have tended to go their separate ways. In the course of our story we will see how these theories come together to explain chess skill; we will see the important constraint that a limited-capacity short-term memory imposes on problem solving in chess and how this limit can be bypassed by specific perceptual knowledge acquired through long experience, stored in long-term memory, and accessed by perceptual discrimination processes.

The phenomena

In Amsterdam, Adriaan de Groot, who was the first psychologist to carry out extensive experiments on problem solving using chess as the task, also initially formulated his theory in terms of heuristic search (*5*). His subjects ranged from quite ordinary players to some of the strongest chess grandmasters in the world, including several former world champions. He was puzzled by one thing: none of the statistics he computed to characterize his subjects' search processes—number of moves examined, depth of search, speed of search—distinguished the grandmasters from the ordinary players. He could only separate them by the fact that the grandmasters usually chose the strongest move in the position, while ordinary players often chose weaker moves. Why were the grandmasters able to do this? Wherein lay their chess skill?

The perceptual basis of chess mastery. One clue to this riddle came when de Groot repeated and extended an experiment that had been performed earlier in the USSR (*6*). He displayed a chess position to his subjects for a very brief period of time (2 to 10 seconds) and then asked them to reconstruct the position from memory. These positions were from actual master games, but games unknown to his subjects. The results were dramatic. Grandmasters and masters were able to reproduce, with almost perfect accuracy (about 93% correct), positions containing about 25 pieces. There was a quite sharp drop-off in performance somewhere near the boundary between players classified as masters, who did nearly as well as grandmasters, and players classified as experts, who did significantly worse (about 72%). Good amateurs (Class A players in the American rating scheme) could replace only about half the pieces in the same positions, and novice players (from our own experiments) could recall only about eight pieces (about 33%). There is a quite nice gradation on this perceptual task as a function of chess skill, and we have verified this in our own experiments (*7*).

We went one step further: we took the same pieces that were used in the previous experiment, but now constructed random positions with them. Under the same conditions, all players, from master to novice, recalled only about three or four pieces on the average—performing significantly more poorly here than the novice did on the real positions. (The same result was obtained by W. Lemmens and R. W. Jongman in the Amsterdam laboratory, but their data have never been published, *8*.)

In sum, these experiments show that chess skill cannot be detected from the gross characteristics of the search processes of chess players but can be detected easily using a perceptual task with meaningful chess content. The experiment with random boards shows that the masters' superior performance in the meaningful task cannot be explained in terms of any general superiority in visual imagery. The perceptual skill is chess-specific. Moreover, a theory of problem solving in chess that does not include perceptual processes cannot be an adequate theory—cannot explain the superior ability of the strong player to choose the right moves.

Eye movements at the chess board. The second set of phenomena we must consider are also perceptual, but of a more recent discovery. Explanations in terms of heuristic search postulate that problem solving, and cognition generally, is a serial, one-thing-at-a-time process. (We are oversimplifying matters to make the issue clear, but the oversimplification will suffice for the present.) Many psychologists have found this postulate implausible and have sought for evidence that the human organism engages in extensive parallel processing (*9*). The intuitive feeling that much information can be "acquired at a glance" argues for a parallel processor. Of course, the correctness of the intuition depends both on the amount of information that can actually be acquired and upon what is meant by a "glance." If a glance means a single eye fixation (lasting anywhere from a fifth of a second to a half-second or longer), then we know that there are high-speed serial processes (e.g. short-term memory search, visual scanning) that operate within this time range (*10*). Thus, it is certainly interesting and relevant to find out how the human eye extracts information from a complex visual display like a chess position and to see whether this extraction process is compatible with the assumptions of the heuristic search theories.

A pair of Russian psychologists, Tichomirov and Poznyanskaya, placed an expert before a chess position with instructions to find the best move, and they observed his eye movements during the first 5 seconds of the task (*11*). The eye movements were inconsistent with the hypothesis that the subject, during these 5 seconds, was searching through a tree of possible moves and their replies.

To describe further what Tichomirov and Poznyanskaya found, we must say a word about how the eye operates. The eye has a central region of high resolution, the fovea (about 1° in radius), surrounded by a periphery of decreasingly lower resolution. Most information about visual patterns is acquired while the fovea is fixated on them; and the eye moves abruptly, in so-called saccadic movements, from one point of fixation to the next. There are at most about four or five saccadic movements per second.

In Tichomirov and Poznyanskaya's record of the first 5 seconds of their subject's eye movements, there were about 20 fixations. Most of these centered on squares of the board occupied by pieces that any chess player would consider to be of importance to the position. There were few fixations at the edges or corners of the board or on empty squares. Moreover, a large number of the saccades moved from one piece to another, where the former piece stood in a "chess" relation—that is, an attack or defense relation—to the latter. For example, the eye would move frequently from a pawn to a Knight that attacked it, or to a Knight that defended it, or from a Queen to a pawn it attacked.

It is important to note that the saccadic movements were not random—therefore, that some information must have been acquired peripherally about the target square before the saccade began. From other evidençe, we know that a strong chess

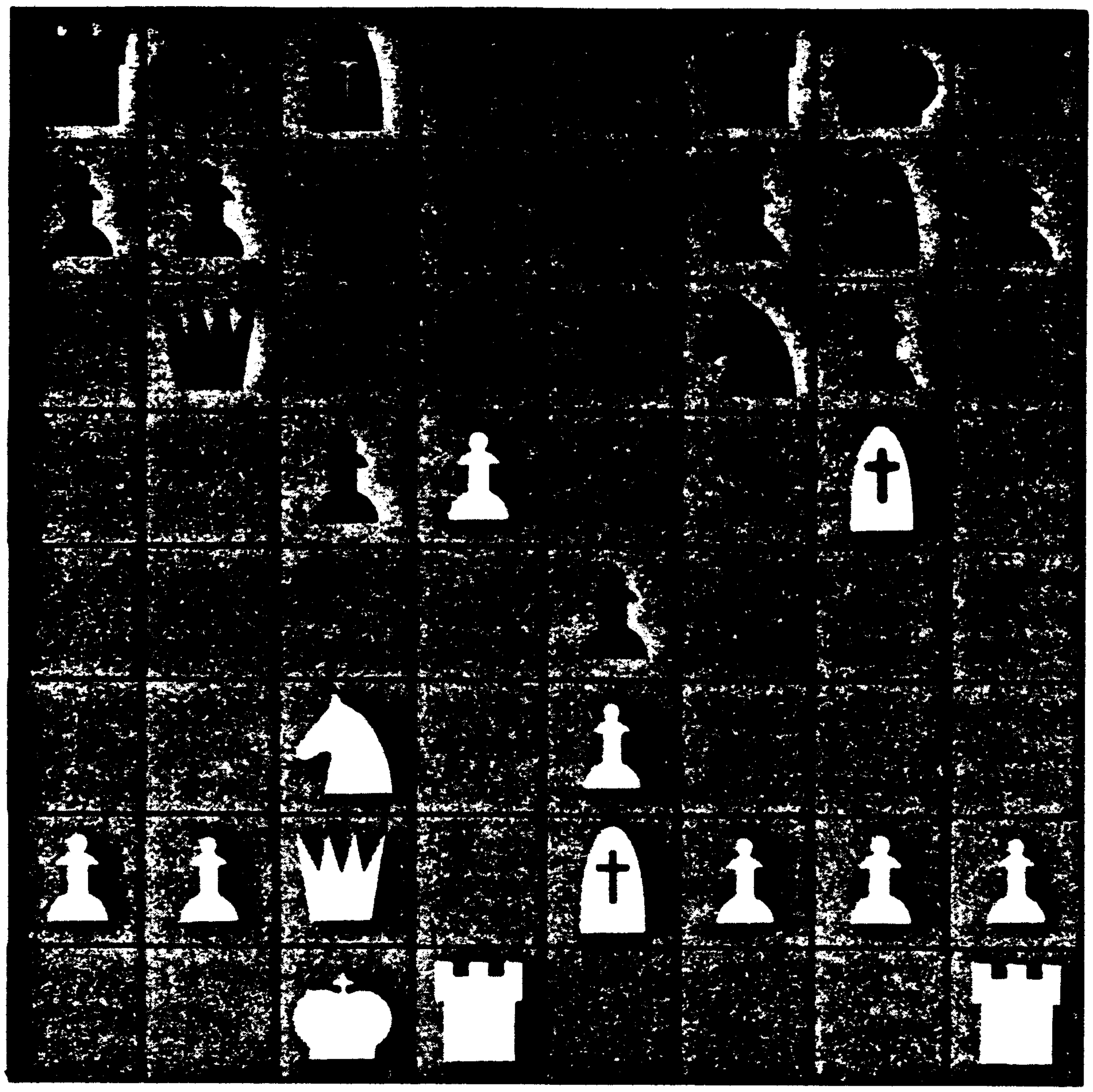

Figure 1. In this middle game position, used by Tichomirov and Poznyanskaya in their eye movement experiments, Black is to play.

player can *recognize* a piece within a radius of 5° to 7° from his point of fixation; for eye-movement studies show that he can frequently replace such a piece correctly on a board when he has had no closer point of fixation to it (*12*).

The Russian experiments are of interest for two reasons. First, while the saccadic eye movements themselves are serial, some parallel visual capacity appears to be operating, for, since the saccade is not random, information about the target square must be acquired peripherally. From what we know about search and scanning rates, it can be concluded that the processes of scanning the periphery for the next target square and preparing the next saccade must overlap in time with the processes of searching memory for the identity and function of a piece (or square) presently occupying the fovea. Visual scanning experiments show that an eye fixation does not allow enough time both to recognize a pattern in the fovea and to scan the visual periphery for a likely target for the next fixation unless the two processes overlap in time (*13, 14*).

Even more important, the Russian experiments confirm the existence of an initial "perceptual phase," earlier hypothesized by de Groot, during which the players first learn the structural patterns of the pieces before they begin to look for a good move in the "search phase" of the problem-solving process. The experiments of Tichomirov and Poznyanskaya have been repeated and confirmed both in Amsterdam and in our own laboratory. How shall we extend the heuristic search theory or problem solving to accommodate them?

Explaining the eye movements

Among the ground rules that ought to be followed in building theories, one of the most important is the rule of parsimony. If, in order to explain each new phenomenon, we must invent a new mechanism, then we have lost the game. Theories, gradually modified and improved over time, are convincing only if the range of phenomena they explain grows more rapidly than the set of mechanisms they postulate.

In the present instance, there are two ways in which we may seek to preserve parsimony as we extend the theory. First, we may examine our existing theory to see whether the mechanisms already incorporated in it might be adequate if they were reorganized. Second, if we need additional mechanisms to explain some of the phenomena, then, instead of inventing them ad hoc, we may draw upon mechanisms already postulated or known in other parts of psychology—mechanisms whose existence already has empirical support. We will explore both of these routes for improving the theory while preserving parsimony.

Perceptual processes in MATER. Let us return to the MATER theory and see how much we must add to, or subtract from, it in order to account for the eye movement data. MATER, as noted earlier, is a program for discovering mating combinations by selective search. What is the basis for the selectivity? A fundamental idea imbedded in MATER is that forceful moves should be explored first, where a forceful move is one that accomplishes some significant chess function, like attacking or capturing a piece or restricting the movements of the opponent. Discovering the opportunities for forceful moves in any chess position involves perceiving the attack, defense, and threat relations that hold among pairs and clusters of pieces on the chessboard—it is basically a perceptual process.

Hence, if we examine MATER a level or two below the executive routine that organizes its search, we see that the program is composed chiefly of a collection of processes for noticing significant chess relations among pieces or squares. In the program as originally organized, these processes were enlisted in the service of the heuristic search for a mating combination. Are these noticing processes a sufficient base on which to build a theory of the eye movements?

The PERCEIVER program. It proved surprisingly easy to simulate the eye movements. It was not difficult to replace MATER's executive program with a new program that used the same perceptual processes to guide the scanning of the board, and when this was done, a good correspondence was found between the squares fixated during the first 20 saccades by the human player and the squares fixated by the program (*15*).

The program, dubbed PERCEIVER, operates in a very simple manner. With the simulated fovea fixated on a square of the board, information is acquired peripherally about pieces standing on nearby squares that attack or defend the fixated square, or that are attacked or defended by the piece on that square. Attention is then assumed to switch to one of these nearby squares, and, unless it immediately returns to the square already fixated, causes a saccadic movement to the new square. With the fovea fixated on the new square, the process simply repeats. A moment's reflection will convince the reader that a process having this structure will cause a biased random walk of the fixation point around the board, returning most frequently to those regions where relations among pieces are densest and spending little time on the edges of the board.

Figure 1 is one of the positions used by Tichomirov and Poznyanskaya in their eye-movement experiments; Figure 2 is a record of the first 20 fixations of their expert in this position; and Figure 3 shows the first 15 fixations produced by PERCEIVER in the same position. Of interest is the fact that the PERCEIVER simulation, by means of its simple mechanism of attending to attack and defense relations, shows the same preoccupation with the important pieces as does the human expert.

There are three points we need to make about this simulation. First, no new mechanisms were invoked; it was sufficient to reorganize the lower-level perceptual mechanisms of MATER. The difference between the behavior of MATER and the

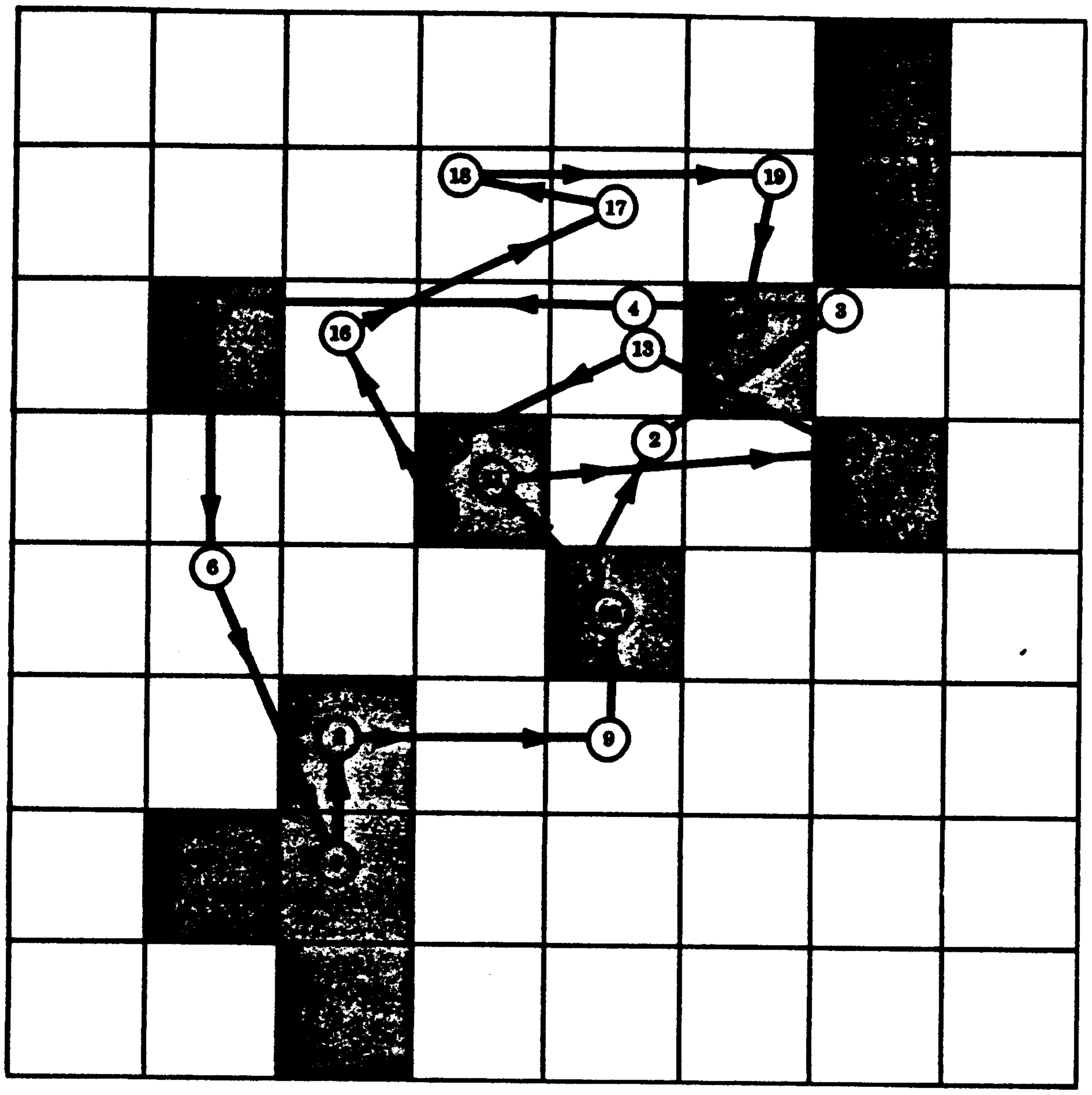

Figure 2. Eye movements of an expert player are recorded for the first 5 seconds, by Tichomirov and Poznyanskaya. The 10 squares occupied by the most active pieces (see Fig. 1) are shaded.

behavior of PERCEIVER lies largely in a difference in goal or motivation at different stages in the problem solving process. The empirical data from human subjects indicate that initially the player sets himself (not necessarily consciously or deliberately, but perhaps habitually) the task of acquiring information about the chess-significant relations on the board (PERCEIVER). Having acquired this information, he turns to generating moves and exploring their consequences (MATER). There would be no great difficulty in revising MATER to conform to this pattern—with the perceptual, information-gathering phase preceding the cognitive, heuristic search phase. As a matter of fact, one earlier computer chess program, written by Newell, Shaw, and Simon in 1958, had much of this flavor (*16*), and another such program is now being constructed by Berliner (*17*).

Second, there is nothing a priori parallel about PERCEIVER; the simple rules that drive the simulated eye around the chessboard are, in fact, serially organized, and it is a simple matter to simulate them in real time on standard computers. Even if realistic time parameters, estimated from human performance, were assigned to the various processes of PERCEIVER, it is still not clear that anything resembling a parallel process would be necessary. This problem is related to the third point.

Third, there is one level of perceptual processing that is finessed and one level that is entirely missing in PERCEIVER. The part that is finessed is the mechanism that recognizes the chess pieces in the first place. What is more important, while PERCEIVER notices attacks and defenses, it has no processes for organizing and remembering this information once it is attended to. But, as we shall see, the organizing process itself drives the eye movements. It is quite plausible that these missing processes operate partly in parallel with the scanning processes of PERCEIVER.

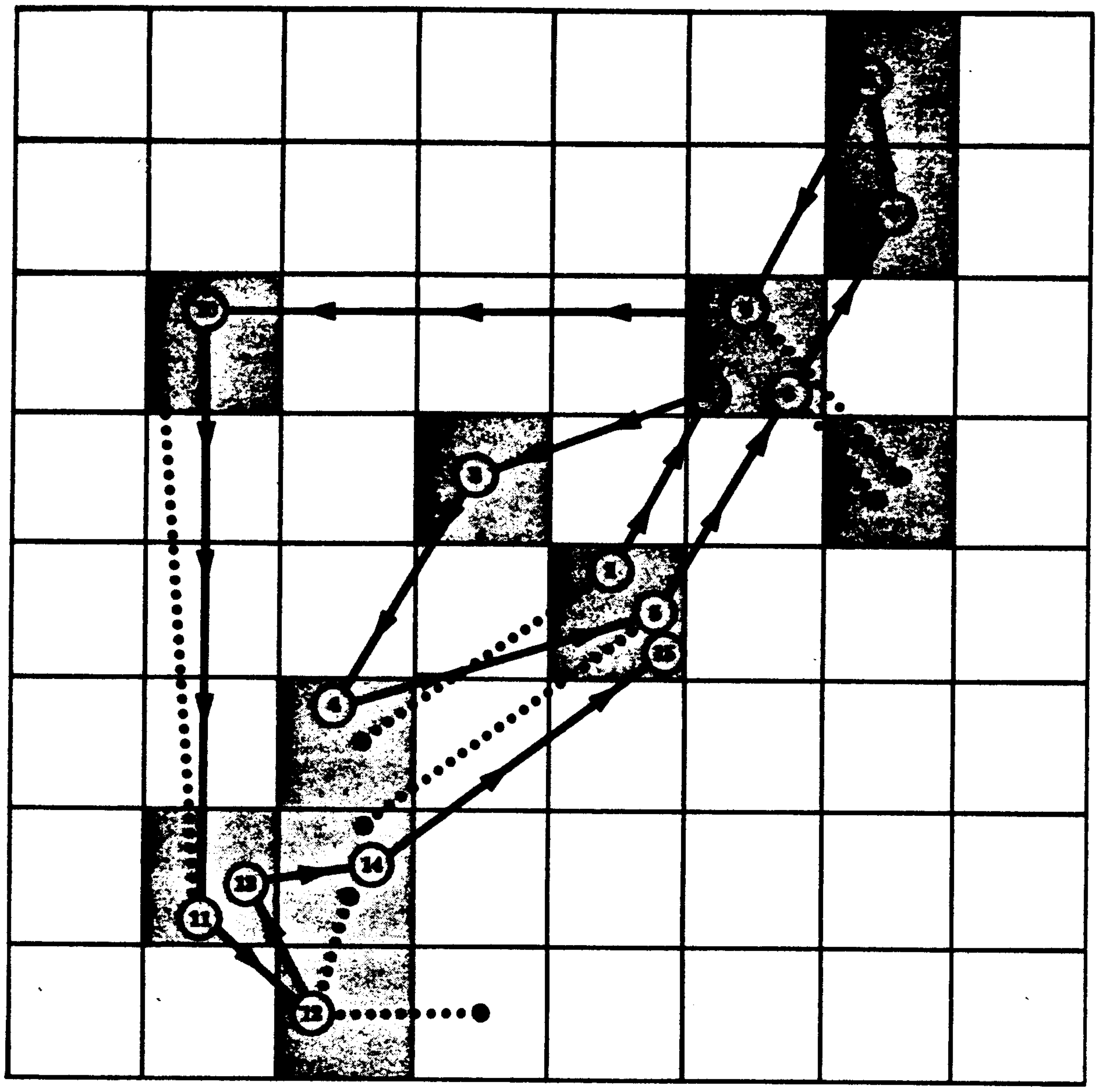

Figure 3. The solid line represents eye movements and the broken lines represent relations noticed peripherally in this record of simulated eye movements during the period of initial orientation from the PERCEIVER program. The 10 squares occupied by the most active pieces (see Fig. 1) are shaded.

Among the striking phenomena that had been observed in rote learning are: (1) a characteristic shape of the serial position curve (in serial anticipation learning), (2) a three-to-one (approximately) time advantage in learning meaningful over meaningless and familiar over unfamiliar syllables, (3) certain characteristic differences in learning times between similar and dissimilar stimulus and response items, and (4) certain conditions that determine whether rote learning will have an incremental or an all-at-once appearance. EPAM has been successful in accounting for all of these phenomena (*19*).

The program of EPAM, and hence the theory it embodies, is quite simple. EPAM learns by growing a discrimination net—a tree-like structure whose nodes contain tests that may be applied to objects that have been described as bundles of perceptual features. When a familiar object is perceived, it is recognized by being sorted through the EPAM net. At the terminal branches of the EPAM net are stored partial "images"—also in the form of feature bundles—of the objects sorted to the respective terminals, together with other information about the objects.

The board reconstruction experiment

Nothing in the perceptual mechanisms we have described so far will allow us to account for the spectacular skill of chess masters in reconstructing positions that they have seen for only a few seconds. Both MATER and PERCEIVER gloss over details of the process for recognizing a chess piece—noticing that it is a Bishop, say, rather than a pawn. Each piece is represented by a little bundle of features—its color, for example, and its type (King, Queen, etc.). The programs do not undertake to explain or simulate the feature extraction process, but simply assume that it is performed and that previous learning has stored in long-term memory the requisite information about the capabilities of the different kinds of pieces. More important, neither program contains any mechanisms for the recognition of meaningful, familiar patterns of pieces—neither program has a mechanism for the extensive storage in long-term memory of familiar patterns, nor indeed do they have a long-term memory of any complexity. But it is precisely this kind of pattern-recognition process that lies at the heart of the master's reconstructive ability.

Elementary perceiver and memorizer. Still retaining our respect for parsimony, we note that there already exists in psychology an information processing theory to explain how feature-bundles can become familiarized, associated with other information in long-term memory, and used as components in larger organizations of structures. This theory, called EPAM (Elementary Perceiver and Memorizer), was initially developed by Feigenbaum to explain some of the principal empirical findings about the rote learning of nonsense syllables in the standard serial anticipation and paired-associate paradigms (*18*).

The EPAM theory also plays an important role in explaining the eye movements. Recall that in the previous section, PERCEIVER was found inadequate because it contained no mechanism for recognizing pieces and patterns of pieces. A more complete theory of eye movements would require that PERCEIVER have access to EPAM.

The processes of EPAM influence the eye movements via the way the discrimination net is searched. Figure 4 illustrates a small section of the net with two terminal nodes. Observe that the nodes contain questions about the contents of specific squares; depending upon what is found at a square, a decision is made concerning which square to query next. In short, the EPAM net is organized as a set of instructions, albeit abstract, for scanning the board for familiar patterns. These instructions must then be interpreted by the perceptual system (PERCEIVER) in order to extract the information, and eye movements may well be necessary to execute the instructions. For small clusters of pieces, some of these successive recognition steps may be executed in a single foveal fixation, without saccadic movement. Thus, eye movements may be of two kinds: (1) initial familiarization, in which simple chess functions (attack, defense) are noticed, and (2) recognition, in which complex patterns are scanned.

This explanation of the eye movements gains additional support from the work of Noton and Stark, who developed independently a similar theory (*20*). They proposed that people's memory of a picture will determine how that picture is subsequently scanned for recognition, and they presented evidence that, under the appropriate conditions, eye movements followed stereotypic "scanpaths" before the picture was recognized. EPAM makes this same strong assumption—that patterns are recognized by scanning the configuration for specific features in a particular order.

EPAM has a recursive structure. This means that any object, once familiarized and incorporated in the net, can itself serve as a perceptual feature of a more complex object. Thus, once the various types of chess pieces—Kings, pawns, Bishops—have become familiarized, these can become features of more complex configurations, say, a "fianchettoed castled Black King's position" (see Fig. 1 for this pattern in the upper-right part of the board). Once familiarized (and this particular pattern is known to every strong player), such a complex can, in turn, serve as a perceptual feature of a still more complex pattern—e.g. an entire chess position.

We have now illustrated the recursive structure of EPAM with a chess example, but the EPAM program was not constructed with this application in mind. In the context of rote verbal learning, the lowest-level features in EPAM are the geometrical and topological properties of English letters. With familiarization, the EPAM net expands to encompass the letters themselves, which then can be used as components (test nodes) of nonsense syllables. Familiarization of the syl-

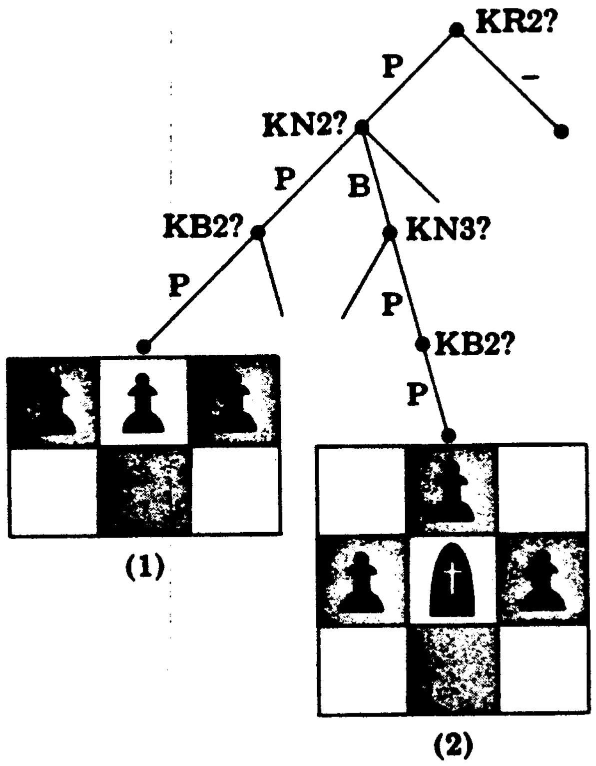

Figure 4. A portion of the EPAM net for chess shows the terminal nodes for two patterns: (1) three pawns on second rank, and (2) fianchettoed Bishop. At each node is shown the test executed there. For KR2?, for example, read: "What piece stands on the King's Rook Two Square?" The patterns at the terminal nodes are for illustrative purposes only: all the information needed to recognize the pattern is imbedded in the logic of the discrimination net. The terminal node has the internal name of the node, an abstract symbolic reference (internal address) that can be stored in short-term memory as a single chunk.

lables, in turn, makes these available as components of syllable pairs or lists, and so on. Thus, EPAM postulates a single learning process, identical with what we have been calling familiarization, and a single kind of output of that process, a new unit or *chunk*.

The EPAM theory implies that the length of time required for a learning task will be proportional to the number of new chunks that have to be familiarized in order to perform the task. This implication also fits the empirical evidence very well, the basic learning time being about 5 seconds per chunk (*21*).

Chunks and short-term memory. Finally, an additional mechanism, short-term memory, is needed in order to understand the reproduction experiment—a mechanism for holding all that information for the short period of time before it is recalled. George Miller, in order to account for the observed invariances in memory-span experiments, first postulated such a memory system with a constant capacity of about seven chunks (*22*). Miller showed that the well-known limit on the amount of information that can be held in short-term memory is not to be measured in bits, but in chunks—the capacity is about "seven, plus or minus two" familiar units of any kind. By acquiring new familiar units (e.g. octal digits) and learning to recode information in terms of those units (e.g. recoding from binary to octal), holding a constant number of *chunks* in short-term memory allows one to hold an increased number of *bits* (in the example, a gain of three to one). The chunk of EPAM theory has these same characteristics.

Since Miller's influential article was published, there has been a tremendous amount of research on short-term memory, and virtually every present-day theory about cognitive processes incorporates such a memory system. Much research on thinking and problem solving has shown that, outside of strategies, the only other human characteristic that consistently limits performance in a wide variety of tasks is the small capacity of short-term memory. And without a short-term memory, EPAM theory by itself does not account for the verbal learning phenomena mentioned earlier. Short-term memory, then, is one of the basic cognitive capacities. For our purposes, we assume that what gets stored in short-term memory are the internal names of chunks (e.g. "fianchettoed castled Black King's position"), which serve as memory addresses or retrieval cues for information about the chunks in long-term memory.

Let us return now to the chessboard construction phenomena. From Miller's chunking hypothesis, EPAM theory, and the limited capacity of short-term memory, we would predict that a chessboard can be reconstructed from information held in short-term memory if, and only if, it can be encoded in not more than about seven familiar perceptual chunks. If a single piece on a particular square constitutes a chunk for a subject, then he should be able to recall only about seven pieces. If he can recall the positions of more than twenty pieces, then it must be that each chunk consists, on average, of a configuration of about three pieces.

We now have a proposed explanation for the remarkable ability of chess masters to reconstruct positions—an explanation that meets our requirements of parsimony. We have employed only mechanisms that are well rooted in other parts of psychological theory: (1) a limited-capacity short-term memory that can hold the names of only about seven chunks, (2) a vast repertoire of familiar patterns stored as chunks in long-term memory, and a recognition mechanism—the EPAM net—for getting at them, and (3) the related chunking process that builds these patterns and their retrieval mechanisms in the first place.

The next task is to find more direct ways to test the theory. Several routes are open: we can seek direct empirical evidence for the existence of these chunks and see if the memory span for chunks is of the order of seven; we can attempt to simulate the reproduction task using the mechanisms of the theory within a computer program; and we can calculate whether the hypothesis leads to reasonable estimates of the number of familiar chunks a chess master must have stored in long-term memory. We consider these in turn.

Empirical identification of chunks. The logic we used in isolating the chunks was to see if, during the reconstruction of a position, chunk boundaries could be identified by long pauses. Time measurements have been used for identifying chunks in other experimental tasks. McLean and Gregg, for example, had subjects memorize permutations of the alphabet (*23*). They then timed the intervals (latencies) between successive letters in the subjects' recitals of the lists. They obtained convincing evidence that the permuted alphabet was stored in memory, not as a single uniform list, but as a hierarchy of segments; the individual letter segments most frequently were three or four letters in length. Within-chunk latencies were much shorter than between-chunk latencies.

Adapting this technique to our task, we videotaped subjects reconstructing chess positions and measured the latencies in placing successive pieces. In order to estimate what interval would correspond to a chunk boundary, we performed a second experiment, in which the subject also reconstructed a chess position but with the original position in view. The two boards were so placed that the subject had to turn his head to look from the one to the other. We found that, when the subject placed two or more pieces on the board without turning his head, each latency was almost always under 2 seconds. We assumed that, under these speeded conditions, subjects load a single chunk into short-term memory when they view the board and then look directly over and recall that chunk. (It would be inefficient, under these conditions, to store more than one chunk, because they would then have to store the chunk names—there isn't enough room in short-term memory to store the structural information comprising more than one chunk—and then at recall use each chunk name in succession to retrieve the chunk from long-term memory—a time-consuming procedure.) We therefore assumed that, in the reconstruction task, a pause longer than 2 seconds indicated the retrieval of a chunk from long-term memory via the chunk name in short-term memory.

To check the plausibility of this 2-second criterion, we counted the number of chess relations that held between pairs of successively placed pieces. The relations counted were attacks, defenses, proximity, identity of type (e.g. both Rooks or pawns), and color. There was a strong negative correlation between numbers of relations and latency (see Fig. 5).

Next, we compared the pattern of frequencies of the between-chunk relations (greater than 2 seconds) with the pattern of the within-chunk relations (less than 2 seconds) and both of these with the pattern that would have been observed had the pieces been replaced in random order. We made this comparison for both forms of the reconstruction experiment—from memory and in sight of the board (see Table 1). For the two forms of the experiment, the within-chunk relational patterns were highly correlated (Pearson correlation coefficient of .89), but these patterns were only slightly correlated with the corresponding between-chunk

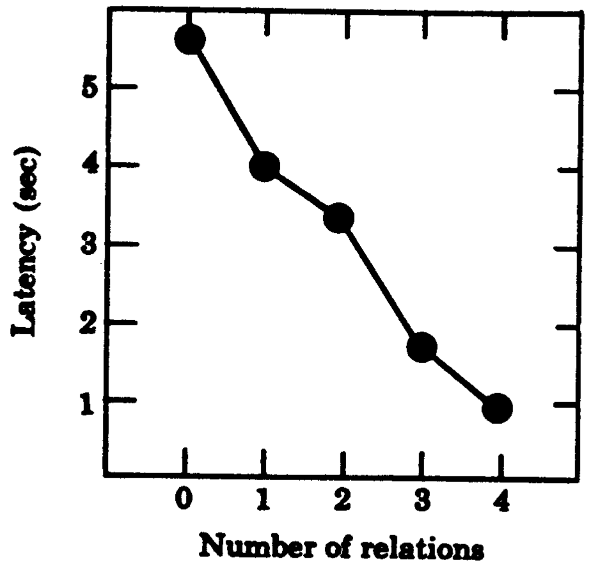

Figure 5. Mean latencies between successively placed pieces in the reconstruction task are plotted as a function of the number of chess relations between the pieces.

patterns (coefficients of .12, .18, .10, and .23) and not at all correlated with the random pattern (−.04 and −.03). On the other hand, the two between-chunk patterns were strongly correlated with each other (.91) and with the random pattern (.87 and .81). Thus, there is strong evidence that the 2-second criterion in fact marks chunk boundaries.

What was the nature of the chunks thus delineated? Most of them were local clusters of pieces in arrangements that recur with high frequency in actual chess positions. (The fianchettoed castled King's position mentioned earlier actually occurs in about ten percent of all recent games between grandmasters.) In the case of a subject who is a chess master, we were able to classify 75% of his chunks as highly stereotyped. Of the 77 chunks observed in his performance of the memory experiment, 47 were pawn chains, sometimes with a nearby supporting or blockading piece. Ten chunks were castled King's positions. Twenty-seven chunks were other clusters of pieces of the same color, and 19 of these were of common types: 9 consisted of pieces on their original squares in the back rank, and 9 of connected Rooks or connected Queen and Rook. These are configurations a chess master has seen thousands of times—as often as we have seen many of the familiar words in our reading vocabularies. There is as much reason to suppose in the one case as in the other that they are stored in his long-term memory and that he will usually recognize them when he sees them.

Table 1. Intercorrelation matrix for the Sight-of-Board Constructions (1 and 3), Memory Constructions (2 and 4), and Hypothetical Random Constructions (5).

	1	2	3	4	5
1. Within-chunk		.89	.12	.18	−.04
2. Less than 2 sec			.10	.23	−.03
3. Between-chunk				.91	.81
4. Greater than 2 sec					.87
5. Random					

Thus far the empirical data support our theory, but we must mention one piece of evidence that is equivocal. If we accept the 2-second criterion for chunk boundaries, then we can measure directly the number of chunks our subjects are holding in short-term memory when they attempt to reconstruct the board. Our theory predicts that the number of chunks will be the same for strong and weak players, but that the average chunk size will vary by a factor of two or three with chess skill.

This prediction is not borne out fully. When we compare, for example, the data from the memory experiment for a chess master with the data for a Class A player, we find that the master recalled about twice as many pieces as the Class A player, but the former's chunks averaged only about 50% larger than the latter's, while the average number of chunks he recalled also averaged about 50% more. The average sizes of the first chunks recalled by master and Class A player were 3.8 and 2.6, respectively; the average numbers of chunks per position were 7.7 and 5.7, respectively. Now the latter numbers are of the right order of magnitude—not far from the memory span of seven—but the difference between them is not predicted by the theory. At the moment, we have no good explanation for the discrepancy, but have simply placed it as an item high on the research agenda. Our hunch is that a less simplistic model of the structure of chunks and their interrelations, or of the organization of chunks in short-term memory, will be needed to attain a better second approximation.

The MAPP simulation. A second approach to testing the theory of the chessboard reconstruction task was to build a computer program, MAPP, to simulate the observed phenomena (*24*). The general outlines of the program follow immediately from our description of the theory. The program contains a learning component to acquire and store in memory a large set of configurations of chess pieces and a performance component to carry out the board reconstruction task (Fig. 6).

Consider first the performance component. When a chess position is presented, the program must scan the board in some way in order to notice the pieces and their relations. The scanning program is a simplified version of PERCEIVER, hence can be viewed as a simulation of the eye movements and control of attention. When a piece is fixated (salient piece), an EPAM-like discrimination process seeks to recognize the cluster of pieces surrounding the fixated piece as a familiar chunk. If it is successful, the symbol designating this chunk is stored in short-term memory. This process is repeated at successive points of fixation until no more pieces become salient or short-term memory capacity is reached, whichever occurs first. Finally, in the reconstruction phase, the terminal information in the EPAM net is used to decode the symbols held in short-term memory into locational information for each of the pieces in a chunk and thus to reconstruct the position.

The learning component of MAPP is a simplified version of the portion of EPAM that grows or elaborates the discrimination net and stores information at its terminal nodes. The input to the program consists of many different configurations of pieces (of two to seven pieces each) that occur frequently as components of chess positions. If such a pattern has been familiarized previously, the program will simply recognize it; if it has not, it will discriminate it from patterns previously learned, will add tests to the EPAM net to implement the discrimination, will create a new terminal node to designate the new pattern, and will store information about the pattern at that node.

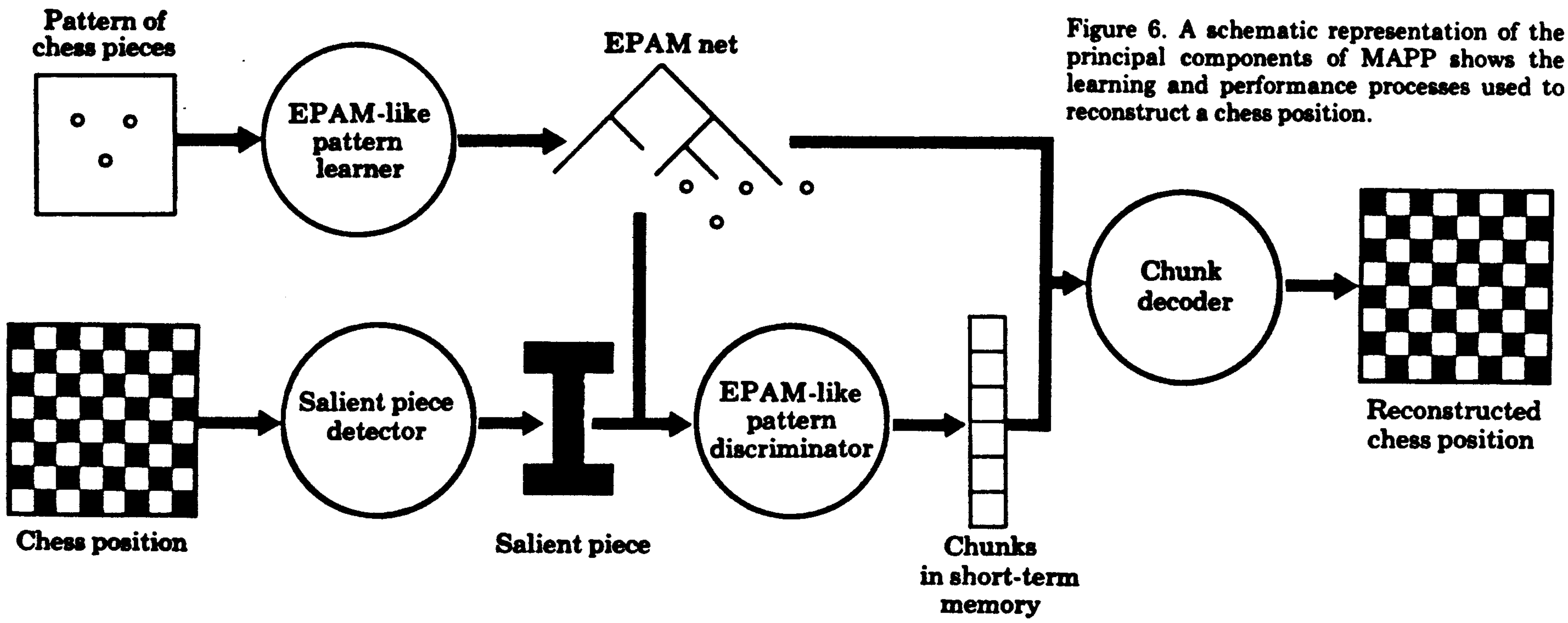

Figure 6. A schematic representation of the principal components of MAPP shows the learning and performance processes used to reconstruct a chess position.

Thus the MAPP program is a hybrid of a simplified PERCEIVER with a simplified EPAM; the finer details of those prior programs are not essential to demonstrating the phenomena. With a net of about 1,000 patterns, the performance of MAPP on the reconstruction task is about equal to that of a Class A player, twice as good as a beginner's, but only half as good as a master's. In a typical set of positions, MAPP recalled 51% of the pieces placed correctly by the master, but only 30% of the pieces missed by the master, indicating that its chunks were not dissimilar from the master's. Finally, the within-chunk chess relations of pieces recalled successively by MAPP were highly similar to those of the human subjects, while the between-chunk relations were close to the random pattern.

The chess master's vocabulary. We can extrapolate from the present performance of the MAPP program to estimate how large a vocabulary of chess patterns would have to be stored in the EPAM net to match the performance of the chess master. The distribution of different patterns by frequency is highly skewed, like the frequency distribution of words in natural language. Assuming that the patterns in the present MAPP net are those most frequently encountered in chess games, and assuming the same degree of skewness for chess patterns as for words, we can estimate that something of the order of 50,000 patterns would have to be stored to match the master's performance. Is this a plausible estimate from other viewpoints? We can check its plausibility in two ways.

First, there are no instant experts in chess—certainly no instant masters or grandmasters. There appears not to be on record any case (including Bobby Fischer) where a person has reached grandmaster level with less than about a decade's intense preoccupation with the game. We would estimate, very roughly, that a master has spent perhaps 10,000 to 50,000 hours staring at chess positions, and a Class A player 1,000 to 5,000 hours. For the master, these times are comparable to the times that highly literate people have spent in reading by the time they reach adulthood. Such people have reading vocabularies of 50,000 words or more. If a chunk is a chunk is a chunk as to learning time (as EPAM theory proposes), then we would expect the chess master to have a comparable chess vocabulary. Our estimate agrees well with that reached previously.

Finally, we may ask: given the variety of possible chess positions from well-played games, how big a vocabulary of patterns must we have so that each position could be represented by a distinct set of seven, or so, patterns? If N is the number of possible positions, while P is the number of patterns, then the requirement is $P^7 > N$. If $P = 50{,}000$, then P^7 is approximately 8×10^{32}. The latter number, in turn, is close to 6^{40}. Now if we played chess games to a depth of 20 moves for each player and at each choice an average of 6 reasonable moves were available, approximately 6^{40} different games could be played. Since there are probably not, on the average, six reasonable moves at each choice point, 50,000 patterns should be more than enough to accommodate the positions that could be reached in such games. It should be emphasized that this estimate is very crude, since it does not take into account that some patterns are much more frequent than others. Nevertheless, it is reassuring that it gives results that are not inconsistent with those arrived at by other routes. Until we can get better data—possibly by expanding the EPAM net—it seems reasonable to assume that a chess master can recognize at least 50,000 different configurations of pieces on sight, and a grandmaster even more.

Familiarity breeds competence

If the MAPP theory provides an explanation—at least a first approximation—of the chess master's superior skill in quickly perceiving chess positions and then reconstructing them from memory, it leaves unexplained the link between this superiority and his chess-playing prowess. How does the theory solve the riddle with which we began—that the statistics of the master's search appear indistinguishable from the statistics of the weaker player's search?

Two facts that have not been much studied in the laboratory, but which are well known in chess circles, need to be mentioned. First, the master and grandmaster not only select good moves but they often—much oftener than weak players—notice these moves in the first few seconds after they look at a new position. Having noticed such a move, the master may continue to analyze the position for some minutes before he is satisfied that it is the best move—and sometimes his analysis will show that his first impulse was wrong. Nevertheless, his ability to notice moves "at a glance" is always astonishing to lesser players.

Second, although the average time per move in serious tournament chess is 3 to 4 minutes (which means that some moves are made rapidly, while others are brooded over for as much as half an hour), a master or grandmaster can beat players of inferior skill while taking only a few seconds per move and playing simultaneously against many players. His play in these games is not of the same quality as in his more deliberate tournament games, but it is strong enough to beat most experts and almost all players of lower class.

The most likely explanation of these facts is that the chess master is not only acquainted with tens of thousands of familiar patterns of pieces, but that with many of these patterns are associated plausible moves that take advantage of the features represented by the pattern (25). Many of the basic heuristics that guide the search for good moves are based on the presence of a pattern on the board. For example, every chess player of even moderate skill is familiar with the advice: "If there's an open file, put a Rook on it." He knows that the advice is not meant quite literally, that what is really meant is "*consider* putting a Rook on it." The pattern of an open file will trigger the heuristic and initiate a move in the heuristic search. Some patterns (perhaps many hundreds) may actually be associated with an algorithmic solution—traps and combinations that lead to the guaranteed win of a piece, a checkmate, or whatnot—in which a series of moves may be played almost by rote.

Thus, we suggest that the key to understanding chess skill—and the solution to our riddle—lies in understanding these perceptual processes. The patterns that masters perceive will suggest good moves to them. The structure of the search process through possible moves will not be very different from that of weaker players; only the paths suggested by the patterns will be different.

Such a view of chess skill is quite amenable to theorizing in terms of production systems. By a *production* is meant a routine consisting of two parts: a *condition* part and an *action* part. The condition part tests the presence or absence of a specific (perceptual) feature (e.g. an open file); the action part, which is executed whenever the condition is satisfied (whenever the feature is recognized as being present), generates a chess move for consideration that is relevant to that specific feature (e.g. putting a Rook on the open file). A separate analysis routine can then carry out the tree search required for a final evaluation of proposed moves. The advantage of modeling human behavior with production systems is that such systems are very simple and rulelike, avoiding many of the inflexibilities of algorithmic programming languages. They can mimic learning by simply adding new productions (*26*), and they have the perceptual flavor we need to simulate the pattern-recognition processes in chess.

While the evidence is not yet in, it becomes increasingly plausible that the cognitive processes underlying skilled chess performance have some such organization as this. Such a scheme would account for the association of chess-playing skill with the ability to recognize numerous perceptual patterns on the board.

There is another question which we haven't addressed directly, but whose answer is implicit in what we have been saying. The question is: how does one become a master in the first place? The answer is *practice*—thousands of hours of practice. This is implicit in the EPAM theory; what is needed is to build up in long-term memory a vast repertoire of patterns and associated plausible moves. Early in practice, these move sequences are arrived at by slow, conscious heuristic search —"If I take that piece, then he takes this piece . . ."—but with practice, the initial condition is seen as a pattern, quickly and unconsciously, and the plausible move comes almost automatically. Such a learning process takes time—years—to build the thousands of familiar chunks needed for master-level chess.

Clearly, practice also interacts with talent, and certain combinations of basic cognitive capacities may have special relevance for chess. But there is no evidence that masters demonstrate more than above-average competence on basic intellectual factors; their talents are chess-specific (although World Champion caliber grandmasters may possess truly exceptional talents along certain dimensions). The acquisition of chess skill depends, in large part, on building up recognition memory for many familiar chess patterns.

We now have an account of perceptual skills in chess that is consistent with theories drawn from other parts of psychology. There is no lack of tasks for continuing research, and the environment of chess continues to be one of the most fruitful for cognitive studies.

References

1. Simon, H. A., and A. Newell. 1964. Information processing in computer and man. *American Scientist* 52:281–300.
2. Simon, H. A., and P. A. Simon. 1962. Trial and error search in solving difficult problems: Evidence from the game of chess. *Behavioral Science* 7:425–29.
3. Baylor, G. W., Jr., and H. A. Simon. 1966. A chess mating combinations program. *AFIPS Conference Proceedings, 1966 Spring Joint Computer Conference* 28:431–47. Washington, D.C.: Spartan Books.
4. Newell, A., and H. A. Simon. 1972. *Human Problem Solving*. Englewood Cliffs, New Jersey: Prentice-Hall.
5. deGroot, A. D. 1965. *Het Denken van den Schaker*. Trans. as *Thought and Choice in Chess*. The Hague: Mouton & Company.
6. Djakow, I. N., N. W. Petrowski, and P. A. Rudik. 1927. *Psychologie des Schachspiel*. Berlin: Walter de Gruyter.
7. Chase, W. G., and H. A. Simon. 1973. Perception in chess. *Cognitive Psychology* 4:55–81.
8. Jongman, R. W. 1968. Het Oog van de Meester. (Doctoral dissertation, University of Amsterdam.) Assen: Van Gorcum & Company.
9. Neisser, U. 1963. The imitation of man by machine. *Science* 139:193–97.
10. Sternberg, S. 1969. Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist* 57:421–57.
11. Tichomirov, O. K., and E. D. Poznyanskaya. 1966. An investigation of visual search as a means of analyzing heuristics. *Soviet Psychology* 5:2–15. (Trans. from *Voprosy Psikhologii* 2(4):39–53).
12. Noordzij, P. 1967. Registratie van oogbewegingen bij schakers. Unpublished working paper, Psychology Laboratory of the University of Amsterdam.
13. Williams, L. G. 1966. The effect of target specification on objects fixated during visual search. *Perception & Psychophysics* 1:315–18.
14. Ellis, S. H., and W. G. Chase. 1971. Parallel processing in item recognition. *Perception & Psychophysics* 10:379–84.
15. Simon, H. A., and M. Barenfeld. 1969. Information-processing analysis of perceptual processes in problem solving. *Psychological Review* 76:473–83.
16. Newell, A., J. C. Shaw, and H. A. Simon. 1958c. Chess-playing programs and the problem of complexity. *IBM Journal of Research and Development* 2:320–35.
17. Private communication.
18. Feigenbaum, E. A. 1961. The simulation of verbal learning behavior. *Proceedings of the Western Joint Computer Conference*, 121–32. (Reprinted in Feigenbaum & Feldman, eds. 1963. *Computers and Thought*. New York: McGraw-Hill.)
19. Simon, H. A., and E. A. Feigenbaum. 1964. An information- processing theory of some effects of similarity, familiarization, and meaningfulness in verbal learning. *Journal of Verbal Learning and Verbal Behavior* 3:385–96.
20. Noton, D. and L. Stark. 1971. Scanpaths in eye movements during pattern perception. *Science* 171:308–11.
21. Simon, H. A. 1969. *The Sciences of the Artificial*. Cambridge: M. I. T. Press, pp. 35–38.
22. Miller, G. A. 1956. The magical number seven, plus or minus two. *Psychological Review* 63:81–97.
23. McLean, R. S., and L. W. Gregg. 1967. Effects of induced chunking on temporal aspects of serial recitation. *Journal of Experimental Psychology* 74(4):455–59.
24. Simon, H. A., and K. Gilmartin. 1973. A simulation of memory for chess positions. *Cognitive Psychology* (in press).
25. Chase, W. G., and H. A. Simon. 1973. The mind's eye in chess. In *Visual Information Processing*, ed. W. G. Chase. Proceedings of Eighth Annual Carnegie Psychology Symposium. New York: Academic Press.
26. Newell, A., and H. A. Simon. 1972. *Human Problem Solving*. Englewood Cliffs, New Jersey: Prentice-Hall.