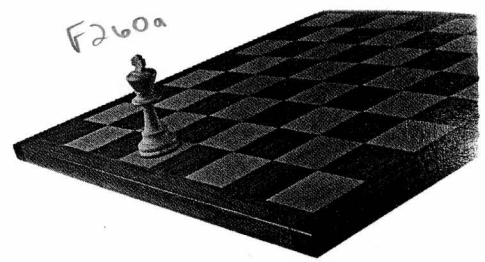# IBM Research

DATA MINING

Decision Support

National Drought Atlas

Inside Deep Blue

LEADERSHIP THROUGH SCIENCE AND TECHNOLOGY

F2600a

THE MAKING OF A
CHESS MACHINE

By Eric J. Lerner

IN BRIEF
How did Deep Blue, IBM's chess-playing computer, manage to compete so vigorously against world chess champion Garry Kasparov last February? Its "brute force" approach was a sophisticated product of the design team's combination of chess, hardware and software expertise. The team now wants to crown its work with a rematch with Kasparov, and also apply the lessons learned on the chessboard to real-life applications.

Designing a chess-playing computer or program that can beat the world champion has been a dream of computer scientists since the early days of digital computers. In 1957, Herbert Simon, computer scientist and later Nobel Prize–winning economist, predicted that a computer would be world champion in a decade. But such predictions underestimated the complexity of the problem. Although the first computer chess programs that could play by the rules emerged in 1958, even the best were playing only at a rank beginner level by 1966, presenting no challenge to an average player.

Initial efforts, guided by the ideas of "artificial intelligence," attempted to have the computer emulate what programmers thought was the way in which human chess players made their decisions. Unfortunately, it became clear that humans use such skills as pattern recognition and associative memory, which have proved extremely difficult for computers to emulate — as well as rather undefined processes, such as intuition and imagination. The lack of success with this approach, coupled with the advent of much faster computers, shifted attention to what IBM team members call the "engineering approach," which takes advantage of a computer's inherent strengths, rather than trying to imitate human thought. This approach relies on the computer's ability to examine large numbers of positions as it tries to determine the best move; the process has, therefore, been labeled as a "brute force" technique.

Simply having a computer look at huge numbers of possible moves and their consequences far into the future runs afoul of chess's complexity. From an average position, there are 38 possible moves. So, looking ahead 6 moves, or 12 "plies" (a ply is a move by one player), involves looking at 38 to the 12th power, or 9 billion billion, positions — far beyond the capabilities of any computer in the foreseeable future. Yet, computers playing at a grandmaster level often require a 6-move look-ahead capability, and sometimes far more. In addition, there is the key question of how to judge in what way one position is better than another: is it having the most pieces, the best threat to the king or the pawns nearest to becoming queens?

By the early 1980s, computer scientists at Bell Laboratories had advanced far enough on these problems to produce a special-purpose chess-playing computer, called Belle, that could search about 4 moves, or 8 plies, forward. That was good enough to play at the United States Chess Federation master level in tournaments against humans. The Bell researchers showed that the faster the machine, the further forward it could look and the better it could play; indeed, it seemed that each additional full move searched could gain 400 rating points. Because Belle played at a chess rating of 2200 (a measure of competitive strength equivalent to a master level), a machine that searched 6 moves, or 12 plies, ahead could, in theory, defeat world champions rated at 2900. However, in practice, the relationship is not linear, and each additional ply beyond the master level contributes a smaller improvement to the rating.

Clearly, far more speed was needed to achieve the depth of search required to reach a grandmaster rating. By the mid-1980s Cray Blitz, running on a Cray supercomputer, and HiTech, running on a special-purpose machine, were searching 100,000 positions per second in the three minutes allowed for the average tournament move.

In 1985, Feng-hsiung Hsu, then a graduate student at Carnegie-Mellon University, designed a very large scale integration (VLSI) chip that could generate 2 million chess moves a second. Hsu and fellow graduate students Thomas Anantharaman, Murray Campbell and Andreas Nowatzyk used spare chips they'd found to put together a chess-playing machine that they called ChipTest. By 1987, the machine, integrating some innovative ideas about search strategies, had become the reigning computer chess champion. A successor, Deep Thought, using two special-purpose chips, plus about 200 off-the-shelf chips, working in parallel, achieved grandmaster-level play.

Shortly after, Hsu, Campbell and Anantharaman (who has since left) joined IBM's Thomas J. Watson Research Center. There, they began work on more powerful chess machines that could beat the world champion, who was then, as now, Russian Garry Kasparov. An intermediate machine, Deep Thought II, incorporated more chess knowledge in the processors and increased the search rate to 3 million to 5 million positions a second. At that pace, it was capable of competing at the grandmaster level.

LIGHTNING SPEED AND HUMAN, RATHER THAN MACHINE, INTELLIGENCE GAVE IBM'S CHESS-PLAYING
COMPUTER THE ABILITY TO COMPETE AGAINST WORLD CHESS CHAMPION GARRY KASPAROV.
THE LESSONS LEARNED IN DESIGNING THE MACHINE MAY HELP FUTURE COMPUTERS TACKLE A VARIETY
OF TASKS, SUCH AS COMPLEX SEARCH PROBLEMS OR AIDING IN THE SYNTHESIS OF NEW DRUGS.

THE ADVENT OF DEEP BLUE

The latest machine, dubbed Deep Blue, incorporates a key change. Instead of having basic chess chips that act merely as move generators, it contains new chips that are complete chess machines, able to generate a position, evaluate it and control the search for further moves. Hsu took nearly three years to design the new chips. The first weren't actually made until September 1995, just five months before the match with Kasparov.

Each chip is a chess powerhouse by itself. Equipped with a million transistors, it can evaluate 2 million positions each second. But in Deep Blue, some 256 chips are teamed together under the overall control of a general-purpose IBM SP2®, a parallel computer consisting of, in this case, 32 processor nodes. The parallelism derived from these 32 processors and 256 chess accelerator chips is what makes Deep Blue the most powerful chess computer in the world. It is capable of looking at an average of 100 million positions per second; that's approximately 20 billion positions in the three minutes allowed for the average tournament move. A general-purpose computer would have to execute a trillion operations per second to match this speed.

Nearly 70 percent of the chip is devoted to evaluating the positions. Here is where the machine's chess knowledge is embedded, making each chip a rather particular sort of "expert system" that encodes human knowledge of a given subject.

Evaluation of a position consists of three parts. A piece-placement evaluation assigns a value for each piece when it is on a given square on the board. "For example," explains Hsu, "a piece might have more value on a more central square, or a pawn on a more advanced square, or a piece might gain value if it's closer to the opponent's King." These values,

downloaded from the software, also change depending on the situation; for a different combination of opening moves, different values would be assigned.

A second evaluation system, used only for the endgame, when few pieces are left, is based in part on stored endgame positions. A third evaluator is the core of the machine's chess knowledge. Here, the chip extracts various features of positions that chess masters have viewed as important and totes up their values. It gives points for pieces that control certain squares (by being

Feng-hsiung Hsu, Joe Hoane and Murray Campbell (*left to right, below*) improved a graduate school project so effectively that, as Deep Blue, it defeated world champion Garry Kasparov in one game, before losing by a respectable score of 4-2.
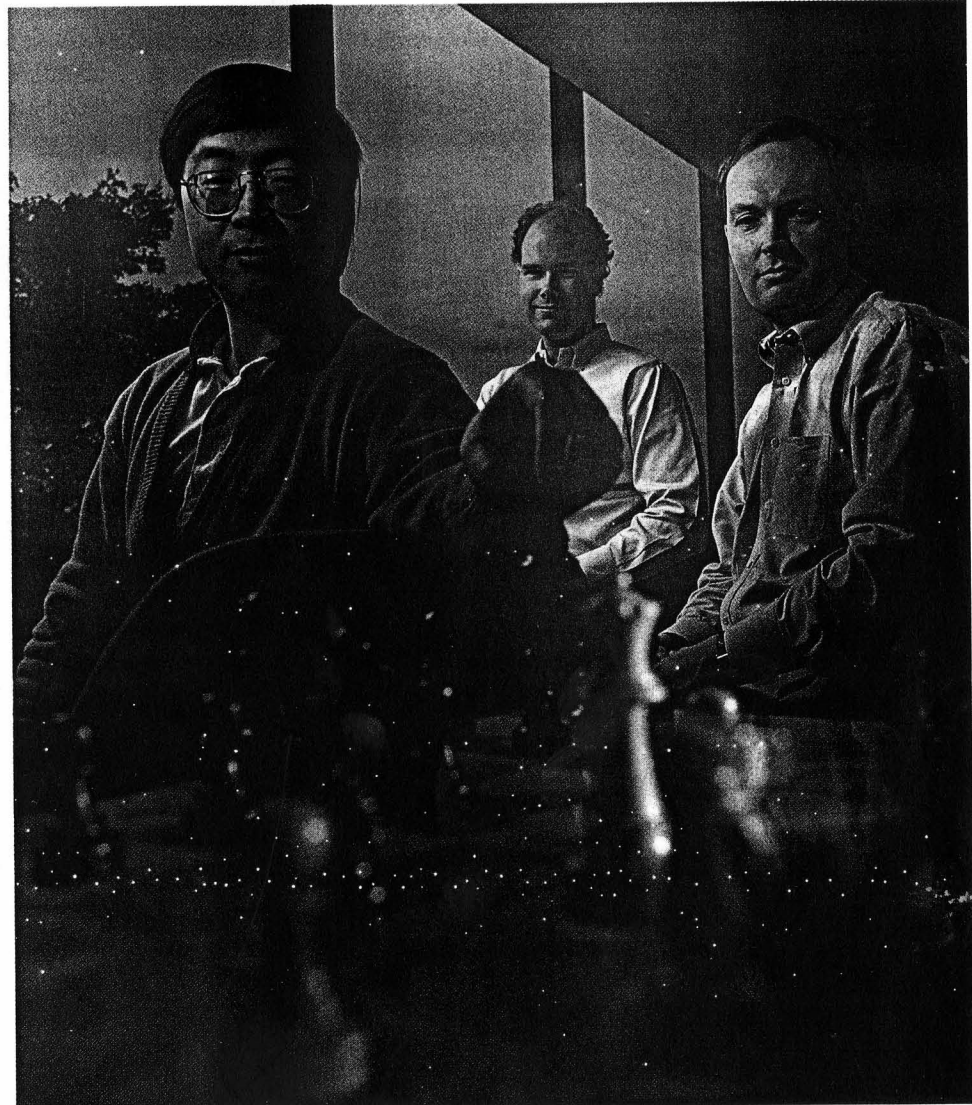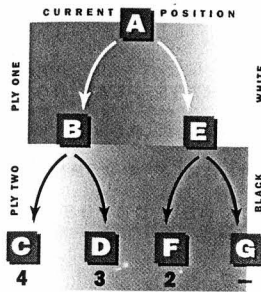
In the alpha-beta search technique, the best move for white is one that leads to the highest possible score, while the best move for black is one that results in the lowest score. Because of those opposed objectives, the computer does not have to evaluate every position. In the example above, the computer — playing white — is looking 2 plies, or 1 move, ahead. It first evaluates the positions that result from moving to B. Depending on black's move, the computer finds that the worst it can do is to obtain a "score" of 3. In evaluating the alternative move that leads to position E, it finds that one of the black's possible moves can force a position with a score of 2. Because that is lower than 3, the computer knows that the move leading to position E would be worse for white than the move leading to B, so it does not have to examine other black moves, such as G.

able to move or capture on those squares); for passed pawns that have gotten past their opponent's pawns; for pins, in which a weaker piece is immobilized because it shields a more valuable one; and so on (*"see Alpha-Beta Search," left*).

The tuning of these features was one of Campbell's main roles in the project, as he is the only strong chess player on the team (his rating of 2150 puts him just below master level). In deciding what features to incorporate, he explains, "We used everything that was important in chess books — we looked at dozens — as well as ideas described in the computer-chess literature."

The consideration of ease also carried weight. "We had to use features that were relatively easy to implement," explains Hsu, "and we invented new ones to try to deal with particular weaknesses we saw in the computer during tournament play." Each feature had to be assigned a value as it applied to each piece or piece combination, for a total of 6,000 values. While the feature extractor and evaluator are wired into the chip, the values are stored in memory on the chip and can be downloaded from the software running on the SP2.

Determining the values is one of the biggest tasks of the team. "We change a value, test it on a set of positions we know the computer has trouble with, test it with games, and then change it again," says Campbell. "This is all done 'by hand' by the programmers, but we are working on automated techniques to tune the evaluation function."

LEARNING FROM MISTAKES

Getting a chess machine to learn from its own mistakes is an appealing idea. It has been tried in the past, but with limited success. "The problem," Campbell explains, "is that when you lose a game, the machine doesn't know what move was the wrong one. It could have been the fourth move or the next-to-last, so it doesn't know what move it has to correct, and determining the reason for the loss and generalizing it to other positions is even more difficult."

In contrast, Deep Blue has no learning ability once its values are chosen by its programmers; it carries out exactly the evaluations hardwired into it. So, in any dictionary definition, as well as in the eyes of its creators, Deep Blue has no intelligence at all. That point seemingly got lost in the media discussion of Deep Blue's IQ.

Even Deep Blue's speed would be inadequate if some way isn't used to narrow the search. In fact, it searches only one billionth of all possible positions in its 6-move look-aheads. Part of the trick, called alpha-beta search, emerged long ago, in the early days of computing. This technique allows the machine to eliminate clearly bad moves from further consideration. In this technique, the machine first looks at a given move, following the possibilities all the way down to 12 plies forward, arriving at a score (say, 103) for the best position that can be expected, given that the opponent plays as well as possible. Then, as it examines alternative moves from a given position, it looks to see if the opponent can force an outcome that is worse (a score lower than 103, in this example). Once it has found that this is the case for a move, it doesn't look at the move further: the move has been "refuted." The machine then goes on to look at the next move until it finds one where there is no refutation, and thus a better outcome can be forced. It continues the search, now looking for refutations that force an outcome lower than this new, better score. This technique allows Deep Blue to cut the average number of moves it has to look at per ply from the 38 moves that are typically allowed to only about 6 moves. This reduces the positions to be examined for 12 plies down to a few billion.

But just searching 6 moves ahead is no good in itself. What if, on the 6th move, the machine has captured an opponent's rook, but, on the 7th move, the opponent is able to force a mate? This "horizon" problem caused early chess machines to frequently go wrong. One of the team's innovations in the earliest versions of Deep Blue was to set up criteria for
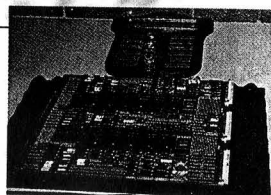

The Path to Deep Blue
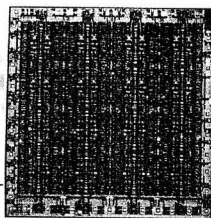
1985: ChipTest
A VLSI chip that could generate 2 million moves per second.

1988: Deep Thought I, containing two specially designed processors, was the first chess computer to beat a grandmaster.

1989: Deep Thought II incorporated more chess knowledge in the processors, greatly increasing its search rate.

much deeper searches on certain positions, called singular extensions. "It means we tend to search deeper when we have fewer choices available, such as in cases where there are checks or piece exchanges," says Joe Hoane, a team member responsible for search software. "This enables Deep Blue to look, in certain situations, up to 30 moves, or 60 plies, ahead, which is what the best players can do."

UP AGAINST KASPAROV

Two months before the match, grandmaster Joel Benjamin joined the team for the first time. His tasks: to help in finding weaknesses in the system and to act as a second during the match. With only one month to go, the last of the 256 chips arrived. The team divided the set into two portions, one with which to play the match and one to serve as backup during the match in the case of equipment failure.

Tests were run up to the last minute. "But we didn't know how well the full-fledged machine would perform until the match actually started," recalls Hsu. While most chess observers expected Kasparov to make short work of the computer, Hsu was cautiously optimistic: "We hoped that we'd make Kasparov sweat," he says.

In fact, the machine ran well enough in the first game to deal Kasparov a stunning defeat, the first victory of a machine over a reigning world champion in regulation play, as opposed to speed matches. Kasparov recovered quickly. With a champion's eye, he began to spot weaknesses in Deep Blue's game. After the the third or fourth game, Kasparov told the audience that "there were some positions the computer played like God and some like 2200" (or barely like a master). Kasparov was able to lead the computer into enough "2200" positions in the second game to defeat it.

Two hard-fought draws followed. Then, early in the fifth game, Kasparov offered another draw. While Deep Blue actually assessed the game as slightly in Kasparov's favor, and thus a good bet for a draw, the IBM team declined the offer. Team members had de-

cided to make decisions on resignations and draws themselves and, in Hsu's words, "We were surprised by such an early offer, and we wanted to see how the game came out." As it happened, Kasparov went on to defeat Deep Blue again and sailed confidently into the final game, and another victory. The final score was 4-2, with three wins for Kasparov, one for Deep Blue and two draws.

THE NEXT MOVES

The Deep Blue project has always had multiple objectives, not all of which are focused on playing championship chess. Nevertheless, says C.J. Tan, the project manager, "The overwhelming public interest in the Kasparov match has encouraged us to explore the possibility of a rematch." If it happens, Kasparov will face an improved opponent across the chess board.

One improvement the team is considering involves making the assignment of values for the evaluator more automatic and scientific, and grandmasters will be employed to help them with the fine-tuning of these values. The team will also modify the system to utilize different styles of play with greater flexibility. Deep Blue will then be able to switch from an aggressive style to a positional style during a match, in the same way a good human player can.

Further matches, however, are not essential to the success of the project. "Basically, we have achieved our goal of building a machine able to play world-champion-level chess," says Tan. "Now, we have initiated activities to use the lessons we have learned to improve the performance of the RS/6000® SP2 to solve other complex problems, such as those that arise in molecular dynamics calculations, drug synthesis and data mining — the extraction of information from massive databases" (*see* "*Digging for Data,*" *page 14*). ■

Eric J. Lerner is a freelance writer based in Lawrenceville, New Jersey.
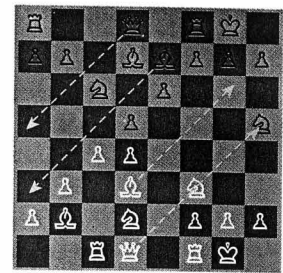
This position, shown after 11 moves of the final game against Garry Kasparov, playing white, provides a simple illustration of the way in which Deep Blue evaluates a position. Deep Blue gives points to Kasparov for such "features" as his white bishop's control over a long diagonal. Similarly, it assigns points to itself, for the control over squares exerted by its black bishop and queen, although those scores are lower than white's because the controlled squares are less central, and hence less valuable. In an additional evaluation scheme, Deep Blue assigns points for the position of each piece on each square. In practice, the software determines Deep Blue's next move by looking at billions of positions that will result from various combinations of moves. It seeks the position that will score highest in the long run, assuming that its opponent, using the same logic, will also try to reach the best position.

1993: Judit Polgar, one of the world's foremost players, lost one game and tied another in her match with the Deep Blue prototype.

WOLF PETER WEBER
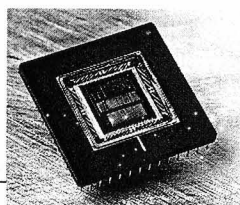
1995: A Deep Blue Chip Under the control of an IBM RS/6000 SP2, some 256 chips form the core of the most powerful chess machine in the world.

One of **Deep Blue's** Two SP2 Frames Its 32 nodes can evaluate 100 million chess positions per second.

1996: Deep Blue vs. Garry Kasparov After a hard-fought match, the final score was 4-2 in favor of Kasparov. Research is readying for a rematch.