

## WHERE IS COMPUTER CHESS GOING?

*Garth E. Courtois Jr.*

Boulder, CO / USA

The Technical Session at the Albuquerque Supercomputing '91 Conference was titled *Where is Computer Chess Going?* The moderator was Tony Marsland, and on the panel were Hans Berliner, Monroe Newborn, Murray Campbell, Robert Levinson, and Jaap van den Herik. Dr. Marsland started the discussion by quoting from a report in the tournament bulletin of remarks made by a similar panel at the 1984 ACM event in San Francisco in response to the question *When will a computer beat the World Chess Champion?* At that time Newborn had estimated, "within 5 years". Ken Thompson said, "eventually". Marsland who had started this discussion by stating, "I once said 'never, or until the limits of humans be known'", now revised that to the year 2020. Then he gave each panel member the opportunity to comment on the question, *Where is computer chess going?* and/or give comments on the World Champion question.

Hans Berliner said, about defeating the World Champion, "Not much longer, perhaps the end of 1994 or 1995." He commented, in the realm of speed-up and knowledge, a few more knowledge items may do it. Good players are not going to take their gloves off for Deep Thought.

Monroe Newborn pointed out that chess will continue on [even after the defeat of the World Champion]. We are just beginning to understand what learning is all about. He expected that programming would be easier but this has not happened (however, he has not given up on it). Computer chess is the auto racing of the computer industry; it may, however, lose followers to computer Go. Another direction is that computers generate their own book. He finished his opening remarks by stating that it is significant that our "Beat the World Champion" estimates are converging.

Jaap van den Herik described the computer world as one with three types of games: there is grandmaster play, machine play, and database / perfect play. It is hard to give these an order of strength for the year 2000. All three have shortcomings. The grandmaster play is understandable, and machines and database play are more "incomprehensible". Grandmasters will study machine play and the machine will try to understand strategy. There will be new research issues with two new final goals. The first is to "crack" the game of chess, the second is to understand the game and strategy.

He then raised the question *What is the error rate in grandmaster play?* He speculated 10%: perhaps in 5 moves (10 half-moves) of a grandmaster game there is one mistake, or 8 mistakes in 40 moves. His second key question to the room was *How do we estimate or measure the error rate of GMs?* There must be some association of rating, playing strength, and understanding. What is the "understanding rating as opposed to the usual rating?" What is the error rate and how do we measure the "understanding rating"? Marsland, as moderator, confronted van den Herik with the question, *How will you get grandmasters to recognize their mistakes? How do you prove there are grandmaster errors?* Dr. van den Herik's answer in essence was to refer to an omniscient database, and he related the story of Karpov-Kasparov (Tilburg, 1991, Round 7). In a six-piece endgame (KBNNKR) which ultimately ended in a draw, the question came up whether they departed from perfect play. Some grandmasters felt the win could be forced because the black King was near a corner. The game was sent to Lewis Stiller<sup>†</sup>, whose computer program had performed a retrograde analysis of this endgame. Analysis of Stiller's reply revealed that these grandmasters made correct moves in terms of departure from the database: neither player had a forced win once this endgame arose. Dr. Berliner's comment along these lines was to reference a different grandmaster game in which Kamsky had a Queen and Pawn against the opponent's Queen. The database showed their position a win in 15 moves. In the course of the game, one player lost a total of 25 moves to per-

† Editors note: see also pages 198-201 in this issue.

fect play, and the other 15 moves. Another noteworthy item was that although they might not have made optimal moves, play was correct in that they did not throw away the theoretical value of the game.

Murray Campbell started his statements with a report on Deep Thought. This version, about one week old, is comparable with the old DT, and they are getting at least one ply better with their 5 million nodes per second. Their hopes are for 10s of millions in late 1992, and 1 billion in early 1993. Dr. Campbell, on *where computer chess is going*, commented that it is important for the computer to play for traps, complex positions which the opponent cannot handle and pointed out that we need to learn how to achieve grandmaster-level evaluation based upon dynamic positions of the tree, and possibly examine pieces in contact. A program without these features will have a difficult time with the world champion. Deep Thought may challenge the world champion by 1994. Marsland posed the question, "Murray, you describe the need for traps and complex positions, yet people associate simplicity with elegance, beauty, and perfection. Is this contradictory?" His response was, "We'll do what's necessary!" Dr. Berliner reinforced this with, "You do what wins."

Robert Levinson feels the first win against a World Champion will be in 1995. A match win will be by the turn of the century, but there will be much human engineering during the match. Computers need to take more responsibility for their evaluation functions. There should also be a better use of grandmaster games if we wish to beat the World Champion.

A member of the audience asked for comments on brute strength vs. technique (intelligent trimming) especially with grandmasters to this 1991 competition. Campbell replied that both approaches will eventually succeed. Larry Kaufman commented from the audience that selective searches outperform full width on comparable hardware. Newborn commented that even a random tree search will eventually succeed. Campbell added that the alpha-beta search is the underlying method in both of these approaches. Marsland said that even in brute-force programs, knowledge enhances the search, such as in move order. Berliner finished with the comment that with a really good evaluator, you need no search!

From the audience, John Stanback asked. "How much work is being done on analyzing grandmaster games and detecting finite tactical errors?" Campbell replied that it is useful to play against the grandmasters later. They are doing this to some extent, especially in the openings. They found a few interesting ideas. Stanback, whose Zarkov book is minimaxed, asked if other books were minimaxed, to which Campbell replied, "Yes."

Don Dailey from the audience mentioned that one main goal is to beat the World Champion. He asked, "How long will it take to cross the extra distance and play a perfect game?" Berliner replied, "Centuries! A perfect game is a science-fiction dream." He struggled with the question of how much better postal-chess game quality compares with over-the-board quality. After considering the possibility of 300-600 rating points, he concluded that we do not know how much better postal chess really is.

Monroe Newborn raised the question on how well a 16-ply program would perform, instead of today's 10 or 11 ply. Berliner replied that there is some doubt a 16-ply program can beat the World Champion. Not just depth is a factor, but knowledge. We do not know that extra plies will scale perfectly to improved performance. 20 ply might help. Van den Herik mentioned his discussions with Karpov stating that 40 ply is certainly enough in the midgame. But what is the minimum? Moreover, there are endgame positions where 40 is not enough. In Van den Herik's opinion, 17 ply is enough to beat the World Champion. Kaufman's comment was that 15 would win a match, but it would not be a convincing victory: perhaps 3-1 with 17 draws. Levinson commented that we have three approaches. The first is brute force, and second, selective search. Number 3 is fundamental mathematical principles, similar to Maxwell's equations. Campbell commented that 1 billion nodes gives us 14 or 15 ply, and added that it is notable that "ply" is becoming a fuzzy measure. There will always be positions where a human can do better so there would not be complete domination over humans.

Robert Levinson of the University of California at Santa Cruz also gave a presentation of the MORPH project, an experience-based creativity program with a potential role of computer chess in the cognitive sciences. MORPH develops from its experiences. It constructs pattern-weight formulations of search knowledge, and these pattern weights can replace subgoals or other search strategies. It uses association retrieval and reinforcement learning. One example of the patterns used helped to specify some of the patterns. Two chess positions were shown with the attack and defense relationships of 5 or 6 pieces. The similarity to the two positions was that White's Queen was attacking and defending the same Knights and Pawns, and all other attack and defenses

were similar, even though the positions looked completely different. The principles learned from one position are used in the other! Another interesting technique mentioned was *simulated annealing*. When MORPH had relatively stable weights, they raised the "temperature", and allowed the weights to cool to a better state.

### Additional Comments

During the tournament I had the opportunity to talk with some of the programmers. Most of the personal computer programs are exploring the selective-search approach. Selectivity is important enough that programs are evolving away from a full-width search, and this departure is serious enough that an "*n*-ply" search no longer clearly describes what the programs are examining. Numbers of ply are reported for the search, and these measures are derived from the Principal Variation. Many of the developers of Zarkov, Socrates, and M-Chess seem to exchange general information among each other. BP is developing in isolation along similar lines. BP, in an *n*-ply iterative-deepening process, does some forward pruning even at ply 1. There are good, interesting, and bad moves. Good ones appear to win material. Interesting ones exchange material, or are the best move in a previous iteration. Bad moves appear to lose material. There is also consideration given to tactical conditions, such as whether there is a piece under attack or if the machine is retreating from check. At ply 1 during early iterations all possible moves are examined. At plies 1 and 2 the good and interesting moves are fully explored, but the bad moves are pruned at *n*-2. Near the bottom of the tree some "serious pruning" is enacted. BP has a 65K position table and would like to make this larger for endgames. It runs on a 33 MHz 486 processor. The iterative deepening goes in steps: 2, 4, ... *n*-2, *n*. When they announce a 6-ply analysis, it is a pseudo-6 ply of full width. They examine approximately 1600 nodes/sec. Part of the BP philosophy is expressed: "If we prune on an even ply, and omit a good move, it is a shame. If we prune on an odd ply and omit a good move, it is a disaster."

### Anecdote

Near the end of the fourth round, the Zarkov - Hitech game was starting to look drawish. John Stanback and Hans Berliner were in good spirits and talking over the board while their computers played. John, realizing that the game was becoming very drawish became concerned about extended play and asked, "We aren't going to have to play until 2 a.m., are we?" Hans replied, "What's your mean time between failures?"



Photo by Terry D. Moore

HOW THEY COMPETE UNDER ACM's AEGIS. (Albuquerque NM, November 1991).

\* Terry D. Moore (9901 Blanda N.W., Albuquerque, NM 87114, USA, 505-898-8232) was ACM's photographer during the 22<sup>nd</sup> ACM tournament, November 1991.