NEWS, INFORMATION, TOURNAMENTS AND REPORTS

MAN AND MACHINE, THEORY AND PRACTICE SQUARE OFF IN SYDNEY

Robert Levinson

University of California at Santa Cruz

1. INTRODUCTION

The 12th International Joint Conference on Artificial Intelligence took place August 24-30 at Darling Harbour in Sydney, Australia – the first time ever in the Southern Hemisphere. Featured on Wednesday, August 28 was a "Chess Afternoon": a 90-minute panel discussion on the Role of Chess in Artificial Intelligence Research, followed by a two-game match between Australian chess champion IM Darryl Johansen and IBM's Deep Thought II. Up to 300 people attended each event, though somewhat fewer for the second match game. Those who attended the afternoon came anticipated great things: several articles had appeared in *The Australian* that week, including one magazine-insert article *Silicon challenges the grandmaster* featuring a picture of Gary Kasparov with the caption "CHECKMATE! – NYET". IBM did a good job of advertising the match at the conference and in the Exhibit Hall, where the Nova Documentary on the match Deep Thought - Kasparov was shown continually.

2. PANEL DISCUSSION

I chaired the panel discussion which included as panelists the computer-chess researchers Tony Marsland and Jonathan Schaeffer, Randy Moulic (representing the Deep-Thought team), and AI planning expert David Wilkins (who is also known for his doctoral work on the PARADISE chess system). Each presentation ran for 15-20 minutes. Moulic described the current state of the Deep-Thought project and showed a video of Feng-hsiung Hsu (who could not attend) speaking on evaluation-function learning in Deep Thought. Moulic also displayed a small chip representing Deep Thought's processor. For a summary of the panel and the positions taken please see *The Role of Chess in Artificial Intelligence Research* at pp. 153-161 of this issue.

The panel was followed by a challenging question from Ed Feigenbaum of Stanford University who wondered why it is that brute-force search is so successful in chess programs but not so in other AI applications, where many believe large quantities of domain knowledge are required to achieve expert performance. My response was, first of all, that it has not yet been demonstrated that brute force is sufficient to beat the World Champion and second that if brute force were applied to other well-structured domains (such as automatic theorem proving or chemical synthesis) with as much seriousness as it has been applied to computer chess, similar successes might be achieved – i.e., now is the time to transfer this research to other domains.

David Wilkins, on the other hand, felt that we were lucky with chess in that it happens to have a certain branching factor and a certain depth for expert play that make it especially well-suited to the high-powered alpha-beta minimax search. He pointed out that the same search-based approach does not yield anything close to expert play in Go or Backgammon because the space that must be searched for championship-level play is so much larger. Thus, chess is not typical even of other board games, much less of important "real-world" applications such as scheduling planes and crews for a major airline.

3. THE CHESS MATCH

After hearing 90 minutes of theoretical and scholarly discussion on AI and computer chess and the relative strengths and weaknesses of man and machine, everybody was eager to see how this manifested on the chessboard. Darryl Johansen (from Melbourne) at 32 is the second strongest player in Australia (behind GM Ian Rogers) and had defeated Deep Thought resoundingly 18 months earlier at the University of Technology, Sydney soon after becoming the Australian champion. With the white pieces, he then played an obscure Re-

verse Sicilian Defense, taking Deep Thought out of its opening book early and grinding it down to win a slow and accurate endgame.

But here Johansen was to face a tougher challenge. Deep Thought II has 10 times the processing power of its predecessor (although, reportedly, it was only using half of its capabilities for this match). It uses 24 specialized processors controlled by an IBM RISC System 6000 computer and examines 7 to 8 million moves a second. The match was scheduled as two games in which each player had 60 minutes for the entire game – a time control that should also make things more difficult for the human. Apparently, Deep Thought II had defeated its predecessor with 2-1 in preparation for the match and had been "studying" Johansen's games over the weekend.

Johansen was playing in a room off-stage under the supervision of Australian arbiter Peter Parr with moves communicated by telephone. Commentary, often humorous and sometimes accurate, was provided on a giant electronic board by IM Terry Shaw, whose analysis failed in Game 1 when Johansen himself was struggling. On being told that the Deep Thought of the future would run on 1000 processors and examine one billion positions, Shaw remarked that he did not even know there were a billion positions! He said that experts had been predicting that a computer would beat the World Champion in ten years for the last thirty years and that in his opinion a computer would beat the champion in about .... ten years! Almost as if to prove this point, Johansen and Deep Thought II (DT II) gave the audience a striking demonstration of their strengths and weaknesses.

3.1. Game 1

In Game 1, with Black, in a tactical duel (the kind Deep Thought thrives on) Johansen blundered early and by move 40 (after giving some harmless checks) was unable to counter each of Deep Thought's threats. Darryl tried to get DT II out of book early (although 3. e4 would have transposed it into a Sicilian) and succeeded only to find that the machine is capable of writing its own book! The move 4. d5! is certainly a surprise, because the white Queen is forced out very early. Darryl may have become overconfident seeing what appears to be a typical (at least in the early days) computer "mistake", because he erred badly with 6. ... Qe7. (The correct move is 6. ... Qf6 since it adds extra protection to the Knight on c6; after 7. Ne5 Nh6 Black appears able to stand his ground although White retains the initiative.) This is the type of tactical mistake a computer (without bugs) would never make. After this, Deep Thought never relinquished its advantage by constantly creating new threats and placing its pieces on powerful squares. By move 12, DT II's advantage looked decisive. With 15. Rd1 and 21. Kd2, DT II lost a bit of time by castling the hard way, using 3 moves to do what a novice can do in one, – maybe it accidentally touched its Rook first? It, perhaps, wasted further time by concentrating almost all its efforts on winning the Pawn on h5. 24. Rh4 looks odd but actually serves two important purposes: it allows the Bishop to go to g4 and prepares the way for doubling Rooks on that file and eventual penetration on the seventh rank. Darryl realized that his only chance is going after the white King with a pawn storm and plays 26. ... d6 in preparation. Unfortunately, this further weakens Black's second rank. The pawn storm resulted in a series of checks exciting the spectators, but Deep Thought has everything under control (probably seeing these checks many moves before) and Darryl was soon forced to resign.

3.2. Game 2

In Game 2, again invoking a Reversed Sicilian (as in the previous game) Johansen playing White produced a dominating and closed center – the type of blocked game where computers, being poor on strategy and long-range planning, have the most difficulty. In wily fashion, Johansen reduced the mobility of each of Deep Thought's pieces and forced resignation by move 35.

By move 6, Johansen already had a strong game, getting d4 in with little contention. 8. ... Bg4 is a serious positional mistake allowing 9. f3 with gain of time and the further bolstering of White's center with e4, reducing Deep Thought to passivity. 16. ... h6 appears to be another positional mistake, weakening the king side. Although White's position is unquestionably strong it is not clear exactly where or how Johansen will cash in. He chose to go after the queen side, where 3 or 4 of DT II's pieces are unable to participate in the proceedings. Of course, most computers, not reasoning schematically, would be unable to appreciate this. After move 19, all of Deep Thought's pieces are on the back two ranks. The seriousness of the bind becomes obvious after 26. ... Qa8, reducing the Queen's mobility to almost nothing. After 31. Qb6, Deep Thought's pieces are draped hopelessly on the edge of the board – the end must be near! The advance of the d-Pawn rams home the last nail in the coffin.

4. CONCLUSIONS

Although these games provide little of interest to the chess theorist (except, perhaps, for Deep Thought II's opening innovation in Game 1), they brought home the critical issue in computer chess very clearly: computers and humans are both capable of dominating the other, though using very different means.

In Game 1 Deep Thought II displayed its tactical strength and played masterfully to maintain the pressure and threats, never losing hold of the tactical initiative. In both games it appeared to display bugs with respect to evaluating king safety (something Deep Thought has had trouble with in the past). In fact, except for not castling, it would be hard to fault its play in any part of Game 1. Even Deep Thought's mistake 8. ... Bg4 in the second game could possibly be attributed to mis-evaluating factors of king safety – seeing 9. f3 as a weak move for White. In Game 2, however, Deep Thought hardly showed up at all (being in a closed positional game) and was soundly beaten in each major aspect of the game: center control, pawn structure, king safety and mobility. Admittedly, Deep Thought II is a very young system and undoubtedly a substantial amount of tuning and debugging will be done in the near future.

Johansen, after his mistake in Game 1, should be given credit, for lasting over 40 moves. In Game 2 by choosing the right opening he was able to steer Deep Thought not only out of its opening book but into exactly the type of game that it has difficulty with. From there, Johansen, displayed the technique ("know-how") that masters have developed from experience, and gave Deep Thought no way out of the bind.

I doubt Deep Thought will ever be more soundly beaten than in Game 2 – without ever achieving an ounce of counterplay. Study of this game may only fuel the fires of those masters and grandmasters who are developing an "anti-computer" style. As computer-chess researchers, we have here been convincingly taught the lesson yet once more – brute force may not be sufficient. Pattern knowledge, planning and strategy may be required as well as the ability to draw on experience. As I stated in the introduction to the panel it is, perhaps, only through understanding and bridging the gap between the way computers and humans play chess that Kasparov can be defeated and real artificial intelligence created. By working to bridge this gap, chess may become a major tool in cognitive-science research. Let us not miss the opportunity!

Deep Thought II - D. Johansen VO 22.6
Sydney (IJCAI'91) Game 1
1. d4 e6 2. Nf3 c5 3. e3 b6 4. d5 exd5 5. Qxd5 Nc6 6. Bc4 Qe7 7. Ne5 Qxe5 8. Qxf7+ Kd8 9. Qxf8+ Kc7 10. Nc3 Nce7 11. Qf3 Kb8 12. e4 Ng6 13. h4 h5 14. Bg5 N8e7 15. Rd1 a6 16. Qe3 Rf8 17. Be2 Nf4 18. Bf3 Nc6 19. g3 Ne6 20. Bxh5 Ncd4 21. Kd2 Nxg5 22. hxg5 Bb7 23. f4 Qe6 24. Rh4 Ka7 25. Bg4 Qf7 26. Rdh1 d6 27. Rh7 b5 28. Kc1 b4 29. Nd5 Bxd5 30. exd5 Rae8 31. Qd3 Qxd5 32. c3 bxc3 33. bxc3 Nb3+ 34. Kc2 Qg2+ 35. Kb1 Qe4 36. Rxg7+ Kb6 37. Rd1 Qxd3+ 38. Rxd3 Re1+ 39. Kc2 Na5 40. Rxd6+ Kb5 41. a4+ Kxa4 42. Bd7+ Ka3 43. Rxa6 Re2+ 44. Kd3 1-0.

D. Johansen - Deep Thought II EO 10.1
Sydney (IJCAI'91) Game 2
1. e3 e5 2. c4 Nf6 3. Nc3 Bb4 4. Nge2 0-0 5. a3 Be7 6. d4 d6 7. d5 c6 8. Ng3 Bg4 9. f3 Bd7 10. Be2 cxd5 11. cxd5 Be8 12. 0-0 Nbd7 13. Kh1 Rc8 14. e4 a6 15. Be3 Kh8 16. Rc1 h6 17. Nf5 Nc5 18. b4 Ncd7 19. a4 Ng8 20. a5 Ngf6 21. Qd2 Rg8 22. Na4 Bf8 23. Nb6 Nxb6 24. Bxb6 Qd7 25. Rxc8 Qxc8 26. Rc1 Qa8 27. Bc7 Nh5 28. Nxd6 f6 29. Qe3 Ba4 30. g3 Be8 31. Qb6 Bf7 32. b5 axb5 33. Bxb5 Be8 34. a6 Bxb5 35. Qxb5 1-0.