

THE HUMANS VERSUS COMPUTERS MATCH 1989

Hans Berliner

School of Computer Science
Carnegie-Mellon University
Pittsburgh, PA 15213 USA

After the sterling performance of the Computer World against the players in the Dutch Championship (cf. *ICCA Journal*, Vol. 12, No. 2, pp. 111-114 and this issue), the results in this event were a big disappointment for computer fans (see the contribution by Dap Hartmann in this issue). Just how drastic their defeat was is the subject of this short note. It is the intention here to focus on the realities of the situation, and to use the results in order to augment our understanding. Part of the cross-table of the five-round Swiss event (including the best-scoring 6 computers only) is presented as Table 1 below.

Computers	Mach IV+	Hitech	Mach IV	Meph. -32	ChipTest	Mach III
Humans						
v.d. Berg	-	1	1	-	1	1
v.d. Vliet	-	-	-	1	1	1
Ree	1	-	0.5	-	1	-
Ligterink	0	1	1	-	-	1
Jongsma	-	0	-	-	-	-
Loewenthal	-	0.5*)	-	-	-	0
Verbaan	0	-	-	-	-	-
Tudjman	-	-	-	-	0	-
Pam	-	0	-	0.5	-	-
Kieboom	0	-	-	-	-	-
Wagenaar	-	-	0	0.5	-	-
Bakker	0.5	-	-	-	-	-
Voorn	-	-	-	0	0	-
v.d. Laan	-	-	-	0.5	-	-
Rek	-	-	-	-	-	-
Voortmeyer	-	-	0	-	-	0
Machine score	3.5	2.5	2.5	2.5	2	2

A

Table 1: The outcome of the best machines versus sixteen human players.

Note: The scores in the boxes are *human* scores, except for the last box.

In order to understand the dimensions of this catastrophe, one must note that there were 7 humans out of 16 who scored better than the best-scoring computer (see elsewhere in this Journal). Then there were 7 more humans out of the remaining 9(!) who scored better than the 2nd highest scoring computer. A complete washout. The 10 computers not mentioned in Table 1 really did not belong in the field and are thus not worthy of analysis. It need only be pointed out in passing that they as a group (all 10) scored only 8.5 point out of 40. The same could

*) The score of Hitech against Loewenthal represents a game that had to be replayed and counted in the tournament table only for Hitech. However, it is part of the data, and should not be ignored.

be said of the humans to the lower side of line "A" with respect to their competitiveness with the top machines, as they lost by a score of 13.5 - 2.5.

All this is very interesting, and is no doubt due to the fact that the organizers wished to get a representative sample of players. However, attempting to stratify the field is very instructive.

Clearly, the top 6 computers could do very well against any but well-versed IM-strength players. First of all, one should congratulate Fidelity Mach IV+ on its outstanding performance. This would be awesome if it were not for the fact that this same type of machine lost to both Deep Thought and Hitech in the World Computer Championships just following. However, this machine played only two of the top 4 humans, but managed an even score. No other machine did nearly as well, although it could be said that its point in the last round against Ligterink was due to the latter trying too hard to win in an attempt to share first place. Nevertheless this was a sterling performance, especially given how well the top 4 humans did otherwise.

So the top 4 humans outscored the top six machines by 12.5 to 1.5, the only machine points coming from two Fidelity products. Why did the machines do so badly against this group? The answer, to this analyst, is that these players were well prepared and played accurately tactically when necessary. It was not unusual to see a human start the game 1. Nc3 d5 2. e4 d4 3. Nce2 and now White has a nice strategic game to play and Black has to invent what to do, which is usually hard for the machines. Only Mach IV (as Black vs. Van den Berg) made much of a fight out of it. Generally, the humans just kept the position simple and steered the game into favorable strategic situations. They demonstrated that this could be done in many ways. I was amazed at how meekly most of these top machines acceded to their opponent's wishes. The humans to the lower side of line "A" apparently were not able to do this.

While the above is a tip of the hat to the top 4 humans, it must be noted that they are not presently considered among the top players in the world. Thus, it would be easy to imagine that those GMs who are now so worried about computer encroachment could learn a lot by a perusal of the games played to the upper side of line "A". (Although it must be said they are principally worried about Deep Thought, and DT did not participate here.)

Finally, I must congratulate Mr. Bakker on his courage and good sportsmanship in coming out to play the monsters. It had been just about some 4 years (December, 1985), since his rather premature prediction that "He would have nothing to fear from computers in the next 5 years", based on a single win versus Cray Blitz [*JCCA Journal*, Vol. 8, No. 4, p. 280]. He played only one of the top six, but did not severely dominate the rest, so I believe one can safely say that he did "have something to fear" and was well advised to turn away my challenge to back up his statement with some money. I hope that next year's event will produce a better showing by the top machines.

THE VALVO-DEEP THOUGHT UNIX MAIL MATCH

Mike Valvo

Towards the end of November, 1988, I proposed a correspondence match between Deep Thought and myself to be conducted over the Netnews facility of the UNIX System. This computer mail system is similar to the LINC except it does not support interactive communication. It contains a newsgroup known as rec.games.chess where it is possible to post an article for all the members of the group to read. Each member is free to post a reply to any article.

The activity on the net was not particularly interesting, so I thought it might be fun to play a game against the computer and post my comments move by move. The computer would not be able to read what I said and everyone could vicariously participate as the game went along.

I had no idea what the results would be. Deep Thought was fresh with victory from the Software Tools tournament and I was near the top of my form in recent years. Deep Thought's USCF rating at the time was 2551 and mine was 2488.