Gobet, F. (1997). Can Deep Blue™ make us happy? Reflections on human and artificial expertise. *AAAI-97 Workshop: Deep Blue vs. Kasparov: The Significance for Artificial Intelligence,* p. 20-23. AAAI Press: Technical Report WS-97-04.

Can Deep Blue™ make us happy? Reflections on human and artificial expertise

Fernand Gobet

ESRC Centre for Research in Development,
Instruction and Training
Department of Psychology
University of Nottingham
University Park
Nottingham NG7 2RD
England
frg@psyc.nott.ac.uk

Abstract

Sadly, progress in AI has confirmed earlier conclusions, reached using formal domains, about the strict limits of human information processing and has also shown that these limits are only partly remedied by intuition. More positively, AI offers mankind a unique avenue to circumvent its cognitive limits: (1) by acting as a prosthesis extending processing capacity and size of the knowledge base; (2) by offering tools for studying our own cognition; and (3) as a consequence of the previous item, by developing tools that increase the quality and quantity of our own thinking. These ideas are illustrated with chess expertise.

> *"...The main thing needed to make the world happy is intelligence. And this, after all, is an optimistic conclusion, because intelligence is a thing that can be fostered by known methods of education."*
>
> Bertrand Russell

Shadows

In itself, the result of the Deep Blue vs. Kasparov match does not matter much (except perhaps to Kasparov). The fact is that almost all chess players are vastly outperformed by Deep Blue and that only a handful of grandmasters can put up any kind of fight against it, as against most chess computers, for that matter. As I will illustrate in this paper, this is but one further episode in the machines vs. humans saga, where humans typically get the worst of it. I will also show that, although humans' egos may be hurt in the process, this outcome is not a reason for gloom. Artificial Intelligence[1] (AI) can help us improve our lives.

The study of human expertise is a hot topic. Top journals in psychology and cognitive science, such *Psychological Review*, *Cognitive Psychology* or *Cognitive Science*, are replete with studies on expert perception, expert memory, and expert problem solving. As more and more becomes known about the psychology of experts, the following conclusion seems inescapable: experts in many domains perform poorly as compared to non-human standards.

It has been known for quite a while that humans perform poorly in highly formal tasks, such as logic (Wason and Johnson-Laird 1976). It has also been known that even experts do not handle basic rules of probability correctly; for example, they constantly misuse base-rate probability (Dawes 1988; Kahneman et al., 1982). Even worse, it has been shown repeatedly, starting with Meehl (1954), that experts in clinical diagnosis are outperformed by simple mathematical predictions tools, such as regression analysis, even when the regression weights are chosen randomly (Dawes 1988). Perhaps, one explanation for these results is that there is a fundamental (architectural) difficulty for us in combining several variables simultaneously.

Chess and intuition

Of course, it is well known that formal techniques do not work in some domains of expertise, either because of their computational complexity or because of the difficulty in fitting the domain into the procrustean bed of mathematics. In such domains, however, humans do manage to reach solutions, perhaps by "intuition" (this unspecified but often proposed mechanism allows, according to its proponents, a holistic understanding of the situation). Chess is a case in point, often used by opponents of classical AI (e.g.,

[1] I will use "Artificial Intelligence" in a broad sense in the paper, including hardware and software, and including some fields that are normally encompassed under computer science, such as data base management.

Dreyfus and Dreyfus 1986; Dreyfus, 1992; Puccetti 1974) and/or defenders of intuition (de Groot 1986) as a domain where symbolic, rule-based thinking techniques are vastly outperformed by human holistic, perceptual thinking. But how good are human solutions in these domains? To answer this question, let us have a look at some empirical results. Again, chess is a good choice, because it is one of the most studied domains in research on expertise.

Research has clearly established that perception plays a key role in human chess expertise. This is reflected in a series of quite amazing abilities that experts have developed in the process of reaching their level, and that differentiate them from novices. For example, experts' eye movements show that they look faster at the key elements of a chess position (de Groot and Gobet 1996). This allows them, as was shown by De Groot, Chase and Simon, to remember positions much better than weaker players (Chase and Simon 1973; de Groot 1965)—they can even recall *random* positions better, though not as well as game positions (Gobet and Simon 1996a). This also allows them to search the problem space very selectively, homing in rapidly on the important variations and pruning many irrelevant branches of the search tree. Empirical evidence for the latter fact comes from rapid-transit chess (Calderwood et al., 1988) and from simultaneous games (Gobet and Simon 1996b). For example, Kasparov plays roughly at his normal strength when simultaneously opposed to eight Masters of international level.

With progress in AI, however, things are turning sour for the proponents of the idea that intuition offers a satisfactory palliative for search: computers are showing us that world experts are not that good even in "preserved" domains such as chess, where mathematical and statistical tools cannot be applied practically. Chess grandmasters and masters are now regularly outperformed by computers. Even worse, comparison with endgame databases show that they play rather poorly in endings that are considered as elementary in textbooks. Consider the ending King-Queen vs. King-Rook, which is typically dealt with in a few pages in endgame textbooks. In a fascinating piece of research, Jansen (1992) has shown that even world-class grandmasters regularly make errors that make winning the game take on average four times longer than the optimal line of play. Because of the presence of the so-called fifty-move rule, this means that, in many cases, they would achieve only a draw instead of a win. (In actual play against a human opponent, they manage to win faster because the defending side makes errors at about the same rate). Interestingly, even authors of textbooks, who have the opportunity to move the pieces on the board and are not subject to time pressure, make errors that would make winning take about twice as long as ideally necessary.

As mentioned above, chess has often been used to illustrate the bankruptcy of rule-based and symbolic thinking and the necessity of supposing a holistic mode of perception. By contrast, the examples just discussed have illustrated situations where these symbolic techniques do better than human intuition, and have even shown where intuition fails. Whether this trend—the victory of rule-based rationality—will be confirmed in the future for other tasks is a fascinating question. For the time being, I will limit myself to addressing the question of what AI can contribute to human cognition.

Lights

Thus, progress in AI has reinforced the conclusions reached earlier by studies of logical and probabilistic reasoning: our capacity for thinking correctly is limited indeed. The human species, although perhaps a bit smarter than other species, is far from having reached a high level of rationality. Gone is the concept that reason is the chief qualify of our species (to quote Russell again, from *Faith and Mountains*: "We think, it is true, but we think so badly that I often feel it would be better if we did not."). Interestingly, one of the main conclusions of the last 30 years of research in AI has been that it is easy to write programs that do better than humans in tasks that tap high-level cognitive functions, but that it is very hard to even approach the low-level perceptual capacities that we share with other mammals.

Is this a reason to despair about mankind's rationality? No, it is not. This is because AI not only shows our limits, but also helps us to overcome them, in several ways.

First, AI augments our processing capacity and the extent of our knowledge base. In chess, use of computers has allowed us to solve endgame questions that had been studied for centuries, and for which quite inaccurate conclusions had sometimes been reached. A typical example is the celebrated endgame King-Queen-g_Pawn vs. King-Queen, where theoreticians (incorrectly) proposed that the best defense was to place the King away from the Pawn. It also allows one to use master game databases to study statistically endgames that are beyond the scope of exhaustive databases and to derive heuristics from the statistical regularities of these endgames (Sturman, 1996; Timoshchenko, 1993). Examples include the relative strengths of Bishop vs. Knight. A final example is offered by Althöfer's experiments showing that a human "collaborating" with a computer produces a game that is superior to that of the human and the computer taken individually (Altöfer, 1997).

In domains more pressing than chess, such as technology and science, similar techniques, including the data mining techniques developed within the Deep Blue technology, make it possible to efficiently and routinely process masses of data the analysis of which was unthinkable twenty years ago. As examples from computational chemistry witness, they also allow us to search huge problem spaces rapidly (Valdes-Perez, 1994). For many of the problems facing us—starvation, pollution, economic recession—problems that contain thousand of variables interacting in complex ways, this extension to our thinking capability probably offers the only hope of finding a solution.

Second, AI offers tools for studying our own cognition. One the one hand, AI offers us standards with which to compare our own performance. There is no doubt that the challenges offered by Deep Blue for chess and Chinook for checkers opened up new dimensions in the play of Kasparov and the late Dr. Tinsley, respectively. In addition, it is hard to over-estimate the impact of computers on the quality of today's play, particularly in our understanding of endgames and openings. Finally, Jansen (1992) has shown how properties of perfect-play databases (in this case, the endgame King-Queen vs. King-Rook) can be studied theoretically and can be manipulated in order to carry out experiments illuminating aspects of human cognition.

On the other hand, AI offers us the mean to construct complex models of chess cognitive processes, as is illustrated in earlier work by Simon (Simon and Barenfeld 1969; Simon and Gilmartin 1973) and in my own work with CHREST (Gobet 1993). These computational models are only a modest illustration of the many computational models, including Soar (Newell 1990) and ACT-R (Anderson 1993) that foster our understanding of cognition.

The presence of computational models and of standards against which to compare human expertise provides powerful tools for cognitive scientists to study our own intelligence. It also offers the possibility of elucidating puzzling concepts such as "intuition," which have eluded understanding because of the lack of adequate experimental environments. With this goal in mind, and given the consent of the human players, parameters in Deep Blue could be varied in order to systematically study factors underlying Grandmasters' famed intuition. Alternatively, Deep Blue parameters could be varied in order to study which constellations produce play that human experts would deem "intuitive." The conclusions reached by these experiments could be used to implement a computational model of intuition.

Third, and perhaps most importantly, AI allows us to use our increased understanding of human cognition to improve our own rationality, by developing software that increases the quality and quantity of our thinking. In chess, artificial tutors are just beginning to make their appearance, offering both instructional material and on-line playing advice. But outside chess, they are already quite common. The best examples of artificial tutors are perhaps offered by Anderson's tutors, based on the ACT-R theory of human cognition, which have been shown to teach skills like LISP programming or geometry more effectively than traditional methods. It is highly probable that the impact of AI on education will be huge in the future (see Anderson et al., 1995, for a review).

Conclusion

It is true that, as AI technology is imported into chess, it also modifies the essence of this game. Chess is a highly competitive domain in which an AI program may be seen either as an unwelcome competitor or as an unfair source of help for the opponent. The presence of computers has already caused the change of several official chess rules and has induced a trend towards more rapid games, where external help is more clearly banned than in traditional 3-minutes-per-move games. Will AI also change the "rules of the game" in domains such as warfare and the economy?

In conclusion, AI is both the bad guy and the good guy: it pinpoints our limits, but it also allows us to go beyond these. As Lord Russell put it so well, intelligence may help us be happy. Had Russell lived until the present time of blossoming AI, I'm sure that he would have said that *Artificial* Intelligence will be a major contributor to the world' happiness.

Acknowledgments

I am grateful to Julian Pine and Frank Ritter for useful comments on this paper.

References

Altöfer, I. (1997). A symbiosis of man and machine beats grandmaster Timoshchenko. *ICCA Journal* 20: 40-47.

Anderson, J. R. (1993). *Rules of the Mind*. Hillsdale, NJ: Lawrence Erlbaum.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences* 4:167-207.

Calderwood, B., Klein, G. A., & Crandall, B. W. (1988). Time pressure, skill, and move quality in chess. *American Journal of Psychology* 101:481-493.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology* 4:55-81.

Dawes, R. M. (1988). *Rational choice in an uncertain world*. Orlando, FL: Harcourt Brace Jovanovich.

de Groot, A. D. (1965). *Thought and choice in chess*. The Hague: Mouton Publishers.

de Groot, A. D. (1986). Intuition in chess. *ICCA Journal* 9:67-75.

de Groot, A. D., & Gobet, F. (1996). *Perception and memory in chess. Heuristics of the professional eye*. Assen: Van Gorcum.

Dreyfus, H. (1992). *What computers still can't do. A critique of artificial reason*. Cambridge, MA: The MIT Press.

Dreyfus, H., & Dreyfus, S. (1986). *Mind over machine.* New York: Free Press.

Gobet, F. (1993). *Les mémoires d'un joueur d'échecs* [The memories of a chess player]. Fribourg (Switzerland): Editions Universitaires.

Gobet, F., & Simon, H. A. (1996a). Recall of rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin & Review* 3:159-163.

Gobet, F., & Simon, H. A. (1996b). The roles of recognition processes and look-ahead search in time-constrained expert problem solving: Evidence from grandmaster level chess. *Psychological Science* 7:52-55.

Jansen, P. J. (1992). *Using knowledge about the opponent in game-tree search.* Ph. D. diss. CMU-CS-92-192, Dept. of Computer Science, Carnegie Mellon, Pittsburgh.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgments under uncertainty: Heuristics and biases.* Cambridge, England: Cambridge University Press.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minneapolis Press.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Puccetti, R. (1974). Pattern recognition in computers and the human brain: With special reference to chess playing machines. *British Journal for the Philosophy of Science* 25:137-154.

Simon, H. A., & Barenfeld, M. (1969). Information processing analysis of perceptual processes in problem solving. *Psychological Review* 7:473-483.

Simon, H. A., & Gilmartin, K. J. (1973). A simulation of memory for chess positions. *Cognitive Psychology* 5:29-46.

Sturman, M. (1996). Beware of the bishop pair. *ICCA Journal* 19:83-93.

Timoshchenko, G. (1993). Bishop or Knight? *ICCA Journal* 16:209-216.

Valdes-Perez, R. E. (1994). Conjecturing hidden entities via simplicity and conservation laws: Machine discovery in chemistry. *Artificial Intelligence* 65:247-280.

Wason, P. C., & Johnson-Laird, P. N. (1976). *Psychology of reasoning. Structure and content.* London: Batsford.