How Intelligent is Deep Blue?

Drew McDermott

[This is the original, long version of an article that appeared in the May 14, 1997 New York Times with more flamboyant title.]

IBM's chess computer, Deep Blue, has shocked the world of chess by defeating Garry Kasparov in a six-game match. It surprised many in computer science as well. Last year, after Kasparov's victory against the previous version, I told the students in my class, ``Introduction to Artificial Intelligence,'' that it would be many years before computers could challenge the best humans. Now that I and many others have been proved wrong, there are a lot of people rushing to assure us that Deep Blue is not actually intelligent, and that its victory this year has no bearing on the future of artificial intelligence as such. I agree that Deep Blue is not actually intelligent, but I think the usual argument for this conclusion is quite faulty, and shows a basic misunderstanding of the goals and methods of artificial intelligence.

Deep Blue is unintelligent because it is so narrow. It can win a chess game, but it can't recognize, much less pick up, a chess piece. It can't even carry on a conversation about the game it just won. Since the essence of intelligence would seem to be breadth, or the ability to react creatively to a wide variety of situations, it's hard to credit Deep Blue with much intelligence.

However, many commentators are insisting that Deep Blue shows no intelligence whatsoever, because it doesn't actually ``understand'' a chess position, but only searches through millions of possible move sequences ``blindly.'' The fallacy in this argument is the assumption that intelligent behavior can only be the result of intelligent cogitation. What the commentators are failing to acknowledge is that if there ever is a truly intelligent computer, then the computations it performs will seem as blind as Deep Blue's. If there is ever a nonvacuous explanation of intelligence, it will explain intelligence by reference to smaller bits of behavior that are not themselves intelligent. Presumably *your brain* works because each of its billions of neurons carry out hundreds of tiny operations per second, none of which in isolation demonstrates any intelligence at all.

When people express the opinion that human grandmasters do not examine 200,000,000 move sequences per second, I ask them, ``How do you know?'' The answer is usually that human grandmasters are not *aware* of searching this number of positions, or *are* aware of searching many fewer. But almost everything that goes on in our minds we are unaware of. I tend to agree that grandmasters are not searching the way Deep Blue does, but whatever they are doing would, if implemented on a computer, seem equally ``blind.'' Suppose most of their skill comes from an ability to compare the current position against 10,000 positions they've studied. (There is some evidence that this is at least partly true.) We call their behavior insightful because they are unaware of the details; the right position among the 10,000 ``just occurs to them.'' If a computer does it, the trick will be revealed; we will see how laboriously it checks the 10,000 positions. Still, if the unconscious version yields intelligent results, and the explicit algorithmic version yields essentially the same results, then they will be intelligent, too.

Another example: Most voice-recognition systems are based on a mathematical theory called Hidden Markov Models. Consider the following argument: ``If a computer recognizes words using Hidden Markov Models, then it doesn't recognize words the way I do. I don't even know what a Hidden Markov Model is. I simply hear the word and it sounds familiar to me.'' I hope this argument sounds silly to you. The truth is that we have no introspective idea how we recognize spoken words. It is perfectly possible that the synaptic connections in our brains are describable, at least approximately, by Hidden Markov Models; if they aren't, then some other equally counterintuitive model is probably valid. Introspection is a lousy way to theorize

about thinking. There are fascinating questions about why we are unaware of so much that goes on in our brains, and why our awareness is the way it is. But we can answer a lot of questions about thinking before we need to answer questions about awareness.

I hope I am not taken as saying that all the problems of artificial intelligence have been solved. I am only pointing out one aspect of what a solution would look like. There are no big breakthroughs on the horizon, no Grand Unified Theory of Thought. Doing better and better at chess has been the result of many small improvements (as was the proof of a novel theorem last year by a computer at Argonne Lab.) There have been other such developments, such as the speech-recognition work I referred to earlier, and many results in computer vision, but few ``breakthroughs.'' As the field has matured, it has focused more and more on incremental progress, while worrying less and less about some magic solution to all the problems of intelligence. A good example is the reaction by AI researchers to neural nets, which are a kind of parallel computer based on ideas from neuroscience. Although the press and some philosophers hailed these as a radical paradigm shift that would solve everything, what has actually happened is that they have been assimilated into the AI toolkit as a technique that appears to work some of the time --- just like Hidden Markov Models, game-tree search, and several other techniques. Of course, there may be some breakthroughs ahead for the field, but it is much more satisfying to get by on a diet of solid but unglamorous results. If we never arrive at a nonvacuous theory of intelligence, we will no doubt uncover a lot of useful theories of more limited mental faculties. And we might as well aim for such a theory.

So, what shall we say about Deep Blue? How about: It's a ``little bit'' intelligent. It knows a tremendous amount about an incredibly narrow area. I have no doubt that Deep Blue's computations differ in detail from a human grandmaster's; but then, human grandmasters differ from each other in many ways. On the other hand, a log of Deep Blue's computations is perfectly intelligible to chess masters; they speak the same language, as it were. That's why the IBM team refused to give game logs to Kasparov during the match; it would be equivalent to bugging the hotel room where he discussed strategy with his seconds. Saying Deep Blue doesn't really think about chess is like saying an airplane doesn't really fly because it doesn't flap its wings.

It's entirely possible that computers will come to seem alive before they come to seem intelligent. The kind of computing power that fuels Deep Blue will also fuel sensors, wheels, and grippers that will allow computers to react physically to things in their environment, including us. They won't seem intelligent, but we may think of them as a weird kind of animal --- one that can play a very good game of chess.