Cogn Comput (2009) 1:17–21
DOI 10.1007/s12559-009-9015-x

Is a Machine Realization of Truly Human-Like Intelligence Achievable?

James L. McClelland

Published online: 18 February 2009

Abstract Even after more than a half a century of research on machine intelligence, humans remain far better than our strongest computing machines at a wide range of natural cognitive tasks, such as object recognition, language comprehension, and planning and acting in contextually appropriate ways. While progress is being made in many of these areas, computers still lack the fluidity, adaptability, open-endedness, creativity, purposefulness, and insightfulness we associate with the supreme achievements of human cognitive ability. Reasons for this and prospects for overcoming these limitations are discussed.

Keywords Human cognition · Open-ended problem solving · Computational theory · Cognitive architecture · Learning algorithms · Nurturance · Culture · Education

Are People Still Smarter than Machines?

In the introductory chapter to *Parallel Distributed Processsing* [18], we began by asking this question:

> Why are people smarter than machines?

At the time, it seemed a good and very important question. The effort to understand and simulate human cognitive abilities had been underway for over three decades, and despite initial promise, seemed not to have gotten very far. To be sure, grand claims had been made. Herbert Simon speaks in his autobiography [20] of announcing to a class in early 1953 that 'Over the Christmas Holidays, Al Newell and I programmed a computer to think'. And the kind of 'thinking' Newell and Simon modeled did produce some impressive results, including the Lisp-based 'Macsyma' [12], a powerful symbolic mathematical system that far exceeded most human's ability to solve mathematical equations. But, as we said on the first page of PDP, the computers and programs of the 1980s were a long way from capturing the fluid, adaptive intelligence people exhibit in a wide range of natural cognitive tasks, including "perceiving objects in natural scenes and noting their relations, understanding language and retrieving appropriate information from memory, making plans, and carrying out contextually appropriate actions."

It is now more than 25 years since these words were written. These years have seen a continuation of the exponential growth in the speed and scale of computers at an ever decreasing price. Desktop computers today are several million times faster and have about 100,000 times more memory than the first commercially available computer (the IBM 704, which went on the market in 1954 [6]), yet at the same time they are also about one thousand times less expensive: The 704 cost two million dollars, and those desktops go for about two thousand. Let us ask, in this context:

> Is it still true that people are smarter than machines?
> And if so: Why?

There's no doubt that there has been progress since the early 80s. For example, in chess, the computer now rules. After some contentious victories and draws in the early years of this decade, a computer chess system, Deep Fritz, beat the undisputed world champion Vladimir Kramnik in 2006. Yet even Deep Fritz did not actually learn to play

J. L. McClelland (✉)
Department of Psychology, Stanford University, Bldg 420, 450 Serra Mall, Stanford, CA 94305, USA
e-mail: mcclelland@stanford.edu

chess; instead some of the smartest minds in the world spent a huge amount of time and money putting together the hardware and software for Deep Fritz, and it could easily be argued that Fritz's play reflects nothing more than clever human programming combined with brute force and table lookup.

What about those 'natural cognitive tasks' we spoke about in the first chapter of *Parallel Distributed Processing*? In vision, computational approaches have made substantial gains. Serre et al. [19] report a neuroscience-inspired feed-forward neural network architecture that learns a general-purpose feature 'dictionary' using an unsupervised learning algorithm. This model is then trained to use the representations learned from unsupervised training to perform an animal/non-animal categorization task, and, after training, matches human performance with brief, masked stimuli. Although the program's training was tailored to the task, the result is still impressive, and further progress seems extremely likely. My impression is that similar progress has occurred in the other natural cognitive tasks we mentioned in our introductory chapter, including language processing and memory retrieval, as well as planning and action selection. While I do not think anyone would claim human-like performance has yet been achieved, I am actually optimistic that incremental progress is occurring in all of these areas, at least up to a point.

Yet, it still seems to me there remain some essential shortfalls in the accomplishments of machine intelligence. To me, a very important limitation is the narrowness of focus one sees in systems that have achieved some degree of artificial intelligence. Consider the computer program I play bridge with on my computer (I play one hand and it plays the other three). The program is pretty good, and yet, there are things it does not take into account that would be taken into account by human players. In a story I once heard, an expert player who we will call Dave has just bid six Hearts, and is about to start play when the director of the club where he is playing announces 'last deal'. Another expert player, Al, from another table that has just finished its last deal, comes over and looks at the hands of all of the players, and lingers to observe the play. The player to Dave's left makes the opening lead. Dave's partner is the dummy. As the dummy lays down his hand, Dave surveys the situation. It looks like an easy contract. But Dave notices that, even after the first trick or two, Al is still hanging around. This makes Dave think: maybe the hand is not such an easy one after all—if it were, Dave would surely have lost interest by now. He ponders: what could conceivably go wrong? Seeing only one possibility—one that would ordinarily seem remote—he devises a plan of play that defends against it, and makes his contract. His opponents are outraged, and complain to the director. But the director can do nothing, since Al never said or did anything.[1]

The story illustrates how humans can bring information from outside a domain to think and reason within it. Few computer programs could do that. Had my computer program been playing Dave's position, it would not have been aware of the presence of the other player and it would not have known how to use that player's presence even if that information was available to it.

This example illustrates a natural characteristic of human thinking: any source of information can play a role in constraining the inferences and plans we make when we make decisions and plan actions. Even the best current cognitive architectures, like ACT-R or SOAR, lack the ability to exploit this kind of situation. Consider the following rule that one might try to write into a computer program in an attempt to overcome this problem: "Always consider whether there is any aspect of the current situation that could provide a hint as to an unanticipated complication." It would be an exciting advance to have a computer program that could evaluate such an open-ended proposition. But this means allowing anything at all to come into play, and appears to leave the computer program in an exhaustive search for all possible inferences of all possible aspects of a situation all of the time. It does not seem likely that this is the way the human mind solves the problem.

Why are People Still Smarter than Machines?

While humans certainly have their shortcomings, the computational approaches that I am familiar with lack the open-ended characteristic of human cognitive abilities illustrated by the example described above, and still depend heavily on the human programmer. I am guessing that few would doubt they also still lack the fluidity, adaptability, creativity, purposefulness, and insightfulness we associate with the supreme achievements of human cognitive ability. While I would not take a strong position of this kind, it seems fairly easy to argue that the real intelligent agent in most artificial intelligence is still the human programmer. Viewed in this way, computers remain, for now, fundamentally nothing more than tools in the hands of their human designers and users, and not autonomous, independent, self-directed, thinking beings, like people.

Why do artificial intelligent systems still have these limitations? One might try to make the case that the problem is still one of sheer computational power. It is

[1] I have heard this story somewhere, but cannot be sure of the source. Perhaps it was in a seminar at MIT in fall of 1982, taught by Jerry Fodor. He certainly would have supported the point that human thought is unlimited in the kind of information it can exploit in reasoning and problem solving.

widely noted that the human brain contains 10^{11} neurons and 10^{15} synapses, and that synapses carry out floating point operations (e.g., multiplying an incoming activation signal times the connection weight) at a temporal resolution approaching about 1000 Hz. That comes to 10^{18} multiplies per second.[2] Both IBM and Sun claim to have broken the petaflop barrier (10^{15} flops per second, CBC News Online, June, 2007 [3]); but that is still three orders of magnitude slower than real time—meaning simulation of 10 min of human cognition would require a full week on such a computer. How quickly supercomputer power actually doubles (and whether the machoflops reported by vendors have any relation to actual performance in real situations) is a matter of debate. If speeds continue to double every 2 years, we should reach the exaflop (10^{18}) level before 2030—so maybe by then we will be able to capture the full scope and scale of human cognitive abilities.

More computer power might be helpful, but it seems pretty clear that this alone will not be sufficient. What other kinds of progress will be necessary? I discuss four that seem most important, drawing on Marr's [11] three well-known levels, but adapting one of them and adding a fourth that is likely to become more and more relevant.

Computational Theory

Marr's three level taxonomy gave cognitive scientists an easy handhold for distinguishing between the fundamental nature and goals of their computational models on the one hand and the algorithms and implementations they use on the other. He also encouraged focus on the computational level itself, something that has continued to gain in importance. The question: "What information is available in the environment, and how can it be optimally used" remains a key question in natural task domains like vision and speech perception. Too often, computer scientists interested in cognitive processes as well as cognitive psychologists interested in computational models have not focused their attention on this question, as Marr so aptly pointed out. And, in spite of considerable progress, we are still a long way from understanding what information is in the stimulus. A simple case in point can perhaps help bring this out. Suppose you see two line segments protruding from either side of an occluder—should you infer that they are connected behind the occluder or not? In the past, researchers investigating this question based their theories on intuitive heuristics that they could turn into equations (e.g., the principle of minimum curvature). More recently, Geissler and Perry [4] have carried out an extensive analysis of the relevant natural scene statistics. Looking at the photographs of natural scenes, they determined the conditional probability that in fact two segments intersecting the same occluder were parts of the same underlying edge as a function of several scene variables. The pattern in these conditional probabilities did not exactly match any of the existing models. And in a follow-up psychophysical experiment, perceiver's judgments matched the scene statistics, not the existing models. This is a tiny example, but one that helps to bring home how much there is to understand about the relationship between stimulus variables and underlying reality. If we are to understand cognitive computation fully, there will be a continuing need to focus on this crucial kind of question.

It should be noted that the issues here are far from trivial. It is very difficult to know exactly how to frame the computational problem. To underscore this point, consider a cognitive system faced with a series of situation–consequence observations in some domain, and let us assume we all agree that it would be a good thing if the system could use the data to learn something about the relationship between situations and consequences. How best should we construe what should be learned in this situation? Currently in the field of cognitive science, there are two views on this question. One holds that we should construe the learner's goal as one of extracting a structured statistical model of the environment—one that explicitly attempts to find the best type of structure to represent the data, and within this the best instance of a structure of a given type [9]. An alternative to this, however, is the position that any taxonomy of alternative types will always provide at best only an approximate characterization of natural structure, so that it is better to define the goal more directly in terms of the problem of optimal prediction, allowing the internal model to remain inchoate instead of explicit (as in a neural network representation; see [17], for further discussion). Both of these ideas require further exploration, and their relations to each other remain to be fully explored. As they currently are construed, the former approach may impose too much constraint, while the latter may impose too little. A deep computational analysis of how constraints can effectively guide the search for optimal solutions to learning problems will surely continue to be an important topic of investigation. There has been some progress on this hard problem (e.g., [23]), but more work is clearly necessary; as things stand, we have guidance on the use of relatively flat solutions to prediction problems, but flat solutions are unlikely to be fully satisfactory, and we have only very small initial steps toward understanding how to guide the search for the right kinds of inchoate multilevel representations.

[2] Only a sub-set of synapses are active during any given millisecond. On the other hand I am leaving out all of the post-synaptic integration, synaptic change, and modulatory influences, not to speak at all of the homeostatic processes continually at play, and so I will stick with the 10^{18} figure as a useful approximation.

Algorithm and Representation

We all need to know what information is in the stimulus and what constitutes the best policy in using it, but that does not mean that we know how a computational mechanism can actually exploit the information effectively. Just what are the best algorithms and representations to use for this purpose?

The debate mentioned above between the more and less structured approaches to define the problem predisposes toward alternative solutions—both of which are computationally intensive. One approach leads to the use of Markov–Chain Monte Carlo search methods, while the other leads to the use of approximate gradient-based approaches like those instantiated in contemporary versions of neural network models, including Deep Belief Networks [2, 5, 16].

To me, an exciting frontier in computational cognitive modeling is the exploration of the computational basis of the characteristics of brain representations, as these have been revealed by recordings from single neurons, and, more recently, from many individual neurons at the same time. There have been exciting developments showing how low-level representations in the visual system [15] and more recently the auditory system [21] can be seen as natural solutions discovered in response to the structure of natural visual and auditory stimuli, and the approach is now being extended to address representations at deeper levels of the processing hierarchy [8].

Architecture

Since fairly early in the days of artificial intelligence, a topic in computational approaches to cognition has been the question of 'the cognitive architecture.' An emphasis on this issue may be one of the primary legacies of the work of Newell and Simon at Carnegie Mellon (see especially [13, 14]). Some such architectures (e.g., ACT-R, [1]) are primarily intended for use in modeling human cognition, while others (e.g., SOAR, [10, 14]) are primarily used as tools for building state-of-the-art artificial cognitive systems. A common theme in the human cognitive modeling literature is to stress some sort of hybrid combination of explicit symbolic and implicit, more connectionist-like, sub-symbolic components [22]. A very recent example of this is SAL, an explicit merger of John Anderson's ACT-R model and Randy O'Reilly's LEABRA architecture [7]. I myself envision a future architecture that is fundamentally sub-symbolic throughout, but which carries out cognitive processes we now envision as symbolic as emergent consequences of the sub-symbolic computations.

One thing that is striking about the approaches described above is that they all rely on the conventional von Neumann computer as the actual underlying computer architecture. Although visionaries have dreamed of fundamentally more parallel and/or brain-like computational systems for quite some time, continual exponential growth in speed and memory capacity have thus far allowed the von Neumann architecture to provide the actual bedrock of most computational models of human cognition, albeit with some degree of multi-processing. However, we may be approaching a singularity is this regard. I have recently been involved in discussions of four very different approaches to a radical reorganization of computation to support truly parallel and interactive processing; and neuromorphic engineering, pioneered by Carver Mead at Caltech over 20 years ago, appears finally to be taking off [24, 26]. It may well be, then, that over the next decade, the butterfly will finally emerge from the chrysalis, and truly parallel computing will take flight.

Nurturance, Culture, and Education

Future improvement in our understanding of the fundamental computational challenges facing cognitive systems, in the algorithms and representations we use to address these challenges, and in the architecture on which these algorithms and representations run are all very likely, and they all seem essential for progress in understanding cognitive computation. Another, additional, step that is needed is to understand the roles of nurturance, culture, and education in structuring human cognitive abilities. Human mental abilities are profoundly shaped by experience, and that experience is structured by social, cultural, and governmental institutions. Even in the first few months of life, when the child is nurtured primarily in the informal social and cultural context of the immediate family, many important changes occur in the child's cognitive, social, emotional, and linguistic capacities that are crucially dependent on the child's experience. The effort to understand how human cognitive abilities arise will depend heavily on taking full account of these influences, and success in achieving true human-like intelligence in artificial systems may rely on the creation of systems that can exploit these influences (see [25]).

References

1. Anderson JR, Lebiere C. The atomic components of thought. Mahwah, NJ: Erlbaum; 1998.
2. Bengio Y, Lamblin P, Popovici D, Larochelle H. Greedy layer-wise training of deep networks. In: Schölkopf B, Platt J, Hoffman T, editors. Advances in neural information processing systems (NIPS). Cambridge, MA: MIT Press; 2007.
3. CBS News Online: IBM and Sun Microsystems hit 'petaflop' speed, June 26, 2007. http://www.cbc.ca/technology/story/2007/06/26/ibm-super.html (2007). Accessed 16 Jan 2009.

4. Geisler WS, Perry JS. Contour statistics in natural images: grouping across occlusions. Visual Neurosci. (in press).
5. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science. 2006;313:504–7.
6. IBM Website: 704 Data processing system. IBM Archives. http://www-03.ibm.com/ibm/history/exhibits/mainframe/mainframe_pp704.html (2008). Accessed 13 Nov 2008.
7. Jilk DJ, Lebiere C, O'Reilly RC, Anderson JR. SAL: an explicitly pluralistic cognitive architecture. J Exp Theor Artif Intell. 2008;20:197–218.
8. Karklin Y, Lewicki MS. Emergence of complex cell properties by learning to generalize in natural scenes. Nature. 2009;457:83–6.
9. Kemp C, Tenenbaum JB. The discovery of structural form. Proc Natl Acad Sci. 2008;105(31):10687–92.
10. Laird JE. Extending the soar cognitive architecture. Memphis, TN: Artificial General Intelligence Conference; 2008.
11. Marr D. Vision. San Francisco: Freeman; 1982.
12. Martin WA, Fateman RJ. The MACSYMA system. In: Proceedings of 2nd symposium on symbolic and algebraic manipulation. New York: ACM; 1971. p. 59–75.
13. Newell A. You cannot play 20 questions with nature and win: projective comments on the papers of this symposium. In: Chase WG, editor. Visual information processing. New York: Academic Press; 1973. p. 283–308.
14. Newell A. Unified theories of cognition. Cambridge, MA: Harvard University Press; 1994.
15. Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature. 1996;381(6583):607–9.
16. Ranzato M, Poultney C, Chopra A, LeCun Y. Efficient learning of sparse representations with an energy-based model. In: Schölkopf B, Platt J, Hoffman T, editors. Advances in neural information processing systems (NIPS). Cambridge, MA: MIT Press; 2007. p. 1137–44.
17. Rogers TT, McClelland JL. Precis of semantic cognition, a parallel distributed processing approach. Behav Brain Sci. 2008;31:689–749.
18. Rumelhart DE, McClelland JL, the PDP research group. Parallel distributed processing: explorations in the microstructure of cognition, Volume I. Cambridge, MA: MIT Press; 1986.
19. Serre T, Oliva A, Poggio T. A feedforward architecture accounts for rapid categorization. Proc Natl Acad Sci. 2008;104(15):6424–9.
20. Simon H. Models of my life. New York: Basic Books; 1991.
21. Smith EC, Lewicki MS. Efficient auditory coding. Nature. 2006;439(7079):978–82.
22. Sun R. Duality of the mind. Mahwah, NJ: Erlbaum; 2002.
23. Vapnik VN. The nature of statistical learning theory. Berlin: Springer; 1995.
24. Vogelstein RJ, Mallik U, Culurciello E, Cauwenberghs G, Etienne-Cummings R. A multi-chip neuromorphic system for spike-based visual information processing. Neural Comput. 2007;19(9):2281–300.
25. Weng J, McClelland JL, Pentland A, Sporns O, Stockman I, Sur M, et al. Autonomous mental development by robots and animals. Science. 2001;291:599–600.
26. Zaghloul KA, Boahen K. A silicon retina that reproduces signals in the optic nerve. J Neural Eng. 2006;3(4):257–67.