

DOI:10.1145/2380656.2380674

Computers interacting with, not imitating, humans is the way forward.

BY ROBERT M. FRENCH

Moving Beyond the Turing Test

ALAN TURING WOULD be 100 years old this year. In 1950 he wrote a seminal paper in which he proposed an operational definition of machine intelligence designed to sidestep the philosophical quagmire of what it means to think.¹⁹ Turing proposed pitting a computer against a human in an “imitation game.” The computer and human are placed in separate rooms and connected by teletype to an external interrogator who can ask any imaginable question of either entity. The computer tries to fool the interrogator into believing it is the human; the human tries to convince the interrogator he or she is the human. If the interrogator cannot distinguish the computer from the person, the computer is judged to be intelligent. This simple test has come to be called the Turing Test.

In the early years of research on artificial intelligence, the test was taken very seriously,^{2,8} especially because many researchers believed truly intelligent machines were just around the corner.^{13,17} But as the 1950s to the 1980s came and went and machines were still no closer to passing the Turing Test, AI researchers began to

realize how difficult the problem of simulating human cognition would actually be.¹² It became clear that human cognition emerges from a complex, tangled web of explicit, knowledge-based processes and automatic, intuitive “subcognitive” processes,¹⁰ the latter deriving largely from humans’ direct interaction with the world. Presumably, by tapping into this subcognitive substrate—something a disembodied computer did not have—a clever interrogator could unflinchingly distinguish a computer from a person.^{5,6} The hope faded that machines would soon be in a position to pass such a test,⁷ and serious researchers in AI focused their energy elsewhere.^{9,20}

In the past decade, however, significant innovation in computer technology and data capture have brought the Turing Test back into focus. That technology, along with vast information resources that became available at the same time, have potentially brought computers closer than ever to passing the test. But, in spite of these developments, the Turing Test still presents significant hurdles, including some unrelated to machine intelligence. Questions based on largely irrelevant aspects of humans’ physiognomy, quirks in their visual, auditory, or tactile systems, and time required to complete various cognitive tasks can be devised to trip up a computer that has not lived life as we humans have with bodies like our own.

In what follows, I argue we need to put aside the attempt to build a

» key insights

- **The time has come to bid farewell to the Turing Test.**
- **Attempting to build a machine to pass a no-holds-barred Turing Test is not the way forward in AI, regardless of recent advances in computing technology.**
- **Building machines that would never pass a Turing Test, but that can interact with humans in a highly meaningful, informative way, is the way forward in AI.**

machine that can flawlessly imitate humans; for example, do we really need to build computers that make spelling mistakes or occasionally add numbers incorrectly, as in Turing's original article,¹⁹ in order to fool people into thinking they are human? So, rather than require a machine to pass a Turing Test and try to proscribe questions that are unfair or inappropriate to judging its intelligence, we should accept the computer as a valid interlocutor and interact with it as an interactive, high-level, sophisticated information source.

Big Data

If we set aside the attempt to build a machine that can pass the Turing Test, can we still make progress in AI that is, nonetheless, in the spirit of the Turing Test? Let us start by considering two fundamental changes that have occurred in recent years—the availability of vast quantities of data of all sorts and the increased speed and power of machines to analyze that data.

The amount of data now available to machines would have been simply unimaginable as recently as 15 years ago; for example, a full visual and auditory recording of 85% of the waking life of an infant from birth to age three exists.^{15,16} Researchers are also developing (and wearing) sophisticated “life-experience” recording devices that allow individuals to record all of the visual and auditory (and potentially, olfactory and tactile) information they experience throughout the day.¹ This means all of the words you might ever utter, hear, read, or write could be stored somewhere as data. Moreover, it could be multiplied by the thousands or even millions of other individuals who also choose to record their lives. To this, add all other available sources of information, from Twitter feeds to Wikipedia, from Facebook to blogs on every conceivable subject, and much, much more.¹⁸ Equally important is the explosion of new algorithms to retrieve, analyze, correlate, and cross-reference this sea of data.

It is reasonable to assume that all of it, appropriately analyzed, would allow a computer to answer questions, including those about people's

reactions to events, as well as their emotions and feelings, it would have had no hope of answering appropriately a decade ago. But there will forever remain “unfair” questions on the Turing Test, such as one from a recent commentary⁴: “Hold up both hands and spread your fingers apart. Now put your palms together and fold your two middle fingers down till the knuckles on both fingers touch each other. While holding this position, one after the other, open and close each pair of opposing fingers by an inch or so. Notice anything?”

Try it yourself. Simply by doing the experiment, you will discover the fact (completely irrelevant as regards intelligence) that you cannot separate your two ring fingers. But how would a computer without a body ever answer the question or a million others like it? And even if, by trolling the Web, someone had reported the answer to this particular body-dependent experiment, there are thousands of other quirky facts, some related to cognitive abilities (such as computation time for multiplication of multi-digit numbers and misspelled words), some with absolutely nothing to do with intelligence, that would trip up a computer. Attempting to define which of these questions is fair or unfair for a Turing Test is not only contrary to the spirit of the Test as originally proposed by Turing but also an endeavor necessarily doomed to failure. My view is: Don't try; accept that machines will not be able to answer them and move on. The point is essentially the same one I made in an earlier essay,⁵ that it would be “essentially impossible for a machine that has not experienced the world as we have to pass the Turing Test.” This observation in no way implies renouncing the goal of building intelligent machines. It suggests merely that we renounce the Turing-inspired goal of building intelligent machines that mimic our own behavior so perfectly that we would not be able to distinguish them from ourselves.

Evolution of Brute Force

The human brain relies on 10^{11} neurons, each with 10^3 synapses, all working in concert to produce cognition. At the lowest level, the brain is



> It seems you
are taking
a certain tone
with me.
An agitated,
human tone...<

>> Really? You
are somewhat
reticent
yourself.<<

> I never said
you were
reticent.<

>> I know.<<

> Smug.<

>> Moi*?<<

> I think
you might be
human.■





> Are you extremely bored answering the same questions?<

>> For all you know, I'm optimized for repetitive tasks! But in all seriousness, it's kinda fun.<<

> If I were you I'd be pretending very hard to be a computer.<

>> please restate the question.
=/ <<

> haha lol etc.■



indisputably performing “mindless” brute-force calculations. In 1997, IBM’s Deep Blue, rated at the time as one of the 300 fastest supercomputers in the world, beat Gary Kasparov, the player generally considered the greatest chess player in history. Deep Blue’s operation was a quintessential example of brute-force search, evaluating some 200 million board positions each second. So, what exactly is the difference between the brute-force computation done by humans and the brute-force computation done by machines? This is a very tricky issue, and there is certainly no simple answer. However, part of the answer involves how brute-force computation evolves.

Imagine that, for some reason, playing high-quality computer chess was essential to human survival. Brute-force search, as practiced by Deep Blue, would likely evolve, by means of automatic techniques akin to genetic algorithms,^{11,14} as well as by explicit human development of ever-more-powerful computer-chess-playing heuristics. People often overlook the fact that Deep Blue’s search of hundreds of millions of board positions per second is inefficient in the extreme, since almost all the positions it considers are completely uninteresting and, therefore, examining them at all is a complete waste of its resources. Consequently, in an evolutionary struggle for survival, such “mindless” brute-force searching would quickly lose out to techniques that channeled brute-force search in ever-more-efficient ways. This is exactly what has happened. Today, there are programs with Elo ratings higher than any human chess player ever and that run on handheld computers. One of the most powerful, Pocket Fritz 4, evaluates “only” 20,000 board positions per second, some four orders of magnitude less than Deep Blue (http://en.wikipedia.org/wiki/Pocket_Fritz).

It is not implausible to imagine this kind of evolution could lead to the emergence in computers of internal representations of board positions and ever-better ways to process these representations. As the internal representations become more complex, better organized in relation to

each other, and processed in ever more sophisticated ways, is it so unreasonable to imagine the gradual emergence of the kind of complexity that would justify the label of a minimal understanding of certain board positions? The bedrock of all understanding is, after all, the ability to construct, contextualize, and make use of internal representations of data.

One of the most impressive recent computer programs to use a combination of brute-force methods and heuristics to achieve human-level cognitive abilities is IBM’s Watson, a 2,880-processor, 80-teraflop computing behemoth with 15 terabytes of RAM that won a “Jeopardy!” challenge in 2011 against two of the best “Jeopardy!” players in history.³ Now imagine that Watson, having beaten the best humans, began to play against programs like itself but that were more computationally efficient than it was. Watson currently has the ability to learn from its mistakes, and, presumably, future algorithms would further improve its search efficiency. Consequently, there is no reason to believe that better and better brute-force computation would not evolve until it had become, like the brute-force computation that underlies our brains, multilayered, hierarchically organized, contextualized, and highly efficient. That is, the brute-force computation of the future will bear as much resemblance to the brute-force algorithms of today as the computers of today resemble the computers of 1950.

What of the Turing Test in all of this? I am convinced no machine will pass a Turing Test, at least not in the foreseeable future, for the overriding reason I outlined earlier: There will remain recondite reaches of human cognition and physiognomy that will be able to serve as the basis for questions used to trip up any machine. So, set the Turing Test aside. I would be perfectly happy if a machine said to me, “Look, I’m a computer, so don’t ask me any questions that require me to have a body to answer, no stuff about what it feels like to fall off a bicycle or have pins and needles in my foot. This fooling you to think I’m a human is passé. I’m not trying to fool you. I’m a computer, ok? On the oth-

er hand, I'd be happy to discuss, say, romantic poetry, the implications of China's one-child policy, or the current financial crisis, and so on. I can't pass a Turing Test, but so what? I can still be a very interesting conversationalist. I don't need to have actually experienced the pain of hitting my hand with a hammer to talk about it. Just like priests don't need to be married to counsel about-to-be-married young people about married life. They know about the trials and tribulations of married life secondhand by having talked about it all their lives to people who are married. So, while their model of marriage might not be as perfect as the model married people have, it is a good enough approximation to provide real insight about marriage. Think of me in those terms."

Computers of the future, even if they never pass a Turing Test, will potentially be able to see patterns and relationships between patterns that we, with all our experience in the world, might simply have missed. The phenomenal computing capacity of computers, along with ever-better data capture, storage, retrieval, and processing algorithms, has given rise to computer programs that play chess, backgammon, Go, and many other highly "cognitive" games as well, or better, than most humans. They compose music and recognize speech, faces, music, smells, and emotions. Not as well as the best humans—not yet—but this is only the beginning. And just as the early failures of AI contributed to our deeper understanding of the true complexity of human cognition, these programs force us to rethink our anthropocentric ideas on the uniqueness of our cognitive skills. But this rethinking should not be a cause for a concern. That a mass of 100 billion slow and imprecise neurons could organize themselves over the course of many millions of years in such a way as to produce human cognition is an amazing outcome of evolution. However, there is no reason to believe this is the only way to achieve cognition. Understanding human cognition and achieving artificial cognition are two separate endeavors, and, even if each can inform the other, they should not be confused.


The goal of building a machine able to pass a Turing Test will long remain elusive and probably never be achieved. But many other great challenges lie ahead, greater even than flawlessly imitating human cognitive behavior down to the last typing mistake. The degree to which progress can be made in AI will be in direct relation to the degree to which the problems to be solved can be represented cleanly and unambiguously. One of the next great challenges of AI will be the development of computer programs designed to discover and prove elegant new mathematical theorems worthy of publication in mathematics journals, not because they were done by a computer but because the mathematics itself will be worthy of publication. Other challenges will be development of programs that make use of the oceans of data now available to find new relationships between diseases and human behavior or the environment. Yet others will be programs that can look at two different pictures and find their analogous elements.

Conclusion

It is time for the Turing Test to take a bow and leave the stage. The way forward in AI does not lie in an attempt to flawlessly simulate human cognition but in trying to design computers capable of developing their own abilities to understand the world and in interacting with these machines in a meaningful manner. Researchers should be clearer about the distinction between using computers to understand human cognition and using them to achieve artificial cognition, meaning we need to revise our long-held notions of "understanding." Understanding is not something only humans are capable of and, as computers get better at representing and contextualizing patterns, making links to other patterns and analyzing these relationships, we will be forced to concede that they, too, are capable of understanding, even if that understanding is not isomorphic to our own. Few people would argue that interacting with people of other cultures does not enrich our own lives and way of looking at the world. In a similar, if not identical way, in the

not-too-distant future, the same will be true of our interactions with computers.

Acknowledgments

This work was funded in part by a grant from the French National Research Agency (ANR-10-065-GETPI-MA). Thanks to Dan Dennett and, especially, to Melanie Mitchell for their insightful comments on an earlier draft of this article. 

References

- Bell, G. and Gemmell, J. *Total Recall: How the E-Memory Revolution Will Change Everything*. Dutton, New York, 2009.
- Dreyfus, H.L. *What Computers Can't Do*. Harper & Row, New York, 1979.
- Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A., Lally, A., Murdock, J., Nyberg, E., Prager, J., Schlaefter, N., and Welty, C. Building Watson: An overview of the DeepQA project. *AI Magazine* 31, 3 (2010), 59–79.
- French, R.M. Dusting off the Turing Test. *Science* 336, 6078 (Apr. 13, 2012), 164–165.
- French, R.M. Subcognition and the limits of the Turing Test. *Mind* 99, 393 (1990), 53–65.
- French, R.M. Subcognitive probing: Hard questions for the Turing Test. In *Proceedings of the 10th Annual Cognitive Science Society Conference* (Montréal, July). LEA, Hillsdale, NJ, 1988, 361–367.
- French, R.M. The Turing Test: The first 50 years. *Trends in Cognitive Sciences* 4, 3 (Mar. 2000), 115–121.
- Haugeland, J. *Artificial Intelligence, The Very Idea*. MIT Press, Cambridge, MA, 1985.
- Hayes, P. and Ford, K. Turing Test considered harmful. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (Montréal). Morgan Kaufman Publishers, San Francisco, 1995, 972–977.
- Hofstadter, D.R. *Metamagical Themas*. Basic Books, New York, 1985, 631–665.
- Holland, J.H. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, 1975.
- Kolata, G. How can computers get common sense? *Science* 217, 4566 (Sept. 1982), 1237–1238.
- Minsky, M. *Computation: Finite and Infinite Machines*. Prentice-Hall, Englewood Cliffs, NJ, 1967.
- Mitchell, M. *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA, 1996.
- Roy, D. Human Speechome Project, 2011; <http://www.media.mit.edu/press/speechome>
- Roy, D., Patel, R., DeCamp, P., Kubat, R., Fleischman, M., Roy, B., Mavridis, N., Tellex, S., Salata, A., Guinness, J., Levit, M., and Gorniaki, P. The Human Speechome Project. In *Proceedings of the 28th Annual Cognitive Science Society Conference*, R. Sun and N. Miyake, Eds. (Vancouver, B.C., Canada). LEA, Hillsdale, NJ, 2006, 2059–2064.
- Simon, H. and Newell, A. Heuristic problem solving: The next advance in operations research. *Operations Research* 6, 1 (Jan.–Feb., 1958), 1–10.
- Talbot, D. A social-media decoder. *Technology Review* (Nov.–Dec. 2011); <http://www.technologyreview.com/computing/38910/>
- Turing, A.M. Computing machinery and intelligence. *Mind* 59, 236 (1950), 433–460.
- Whitby, B. The Turing Test: AI's biggest blind alley? In *Machines and Thought: The Legacy of Alan Turing*, P. Millican and A. Clark, Eds. Oxford University Press, Oxford, U.K., 1996, 53–63.

Robert M. French (robert.french@u-bourgogne.fr) is a Research Director of the French National Center for Scientific Research (CNRS) working at the Laboratoire d'Etude de l'Apprentissage et du Développement at the University of Burgundy, Dijon, France.