Connecting the Humanities and the Sciences: Part 2. Two Schools of Thought: The Turing Approach vs. The Lovelace Approach*

Walter Isaacson, The Jefferson Lecture, National Endowment for the Humanities, May 12, 2014

That brings us to another historical figure, not nearly as famous, but perhaps she should be: Ada Byron, the Countess of Lovelace, often credited with being, in the 1840s, the first computer programmer.

The only legitimate child of the poet Lord Byron, Ada inherited her father's romantic spirit, a trait that her mother tried to temper by having her tutored in math, as if it were an antidote to poetic imagination. When Ada, at age five, showed a preference for geography, Lady Byron ordered that the subject be replaced by additional arithmetic lessons, and her governess soon proudly reported, "she adds up sums of five or six rows of figures with accuracy."

Despite these efforts, Ada developed some of her father's propensities. She had an affair as a young teenager with one of her tutors, and when they were caught and the tutor banished, Ada tried to run away from home to be with him. She was a romantic as well as a rationalist.

The resulting combination produced in Ada a love for what she took to calling "poetical science," which linked her rebellious imagination to an enchantment with numbers.

For many people, including her father, the rarefied sensibilities of the Romantic Era clashed with the technological excitement of the Industrial Revolution. Lord Byron was a Luddite. Seriously. In his maiden and only speech to the House of Lords, he defended the followers of Nedd Ludd who were rampaging against mechanical weaving machines that were putting artisans out of work. But his daughter Ada loved how punch cards instructed those looms to weave beautiful patterns, and she envisioned how this wondrous combination of art and technology could someday be manifest in computers.

Ada's great strength was her ability to appreciate the beauty of mathematics, something that eludes many people, including some who fancy themselves intellectual. She realized that math was a lovely language, one that describes the harmonies of the universe, and it could be poetic at times.

She became friends with Charles Babbage, a British gentleman-inventor who dreamed up a calculating machine called the Analytical Engine. To give it instructions, he adopted the punch cards that were being used by the looms.

Ada's love of both poetry and math primed her to see "great beauty" in such a machine. She wrote a set of notes that showed how it could be programmed to do a variety of tasks. One example she chose was how to generate Bernoulli numbers. I've explained special relativity already, so I'm not going to take on the task of also explaining Bernoulli numbers except to say that they are an exceedingly complex infinite series that plays a role in number theory. Ada wrote charts for a step by step program, complete with subroutines, to generate such numbers, which is what earned her the title of first programmer.

In her notes Ada propounded two concepts of historic significance.

The first was that a programmable machine like the Analytical Engine could do more than just math. Such machines could process not only numbers but anything that could be notated in symbols, such as words or music or graphical displays. In short, she envisioned what we call a computer.

Her second significant concept was that no matter how versatile a machine became, it still would not be able to *think*. "The Analytical Engine has no pretensions whatever to originate anything," she wrote. "It can do whatever we know how to order it to perform… but it has no power of anticipating any analytical relations or truths." [i]

In other words, humans would supply the creativity.

This was in 1842. Flash forward one century.

Alan Turing was a brilliant and tragic English mathematician who helped build the computers that broke the German codes during World War II. He likewise came up with two concepts, both related to those of Lovelace.

The first was a formal description of a universal machine that could perform any logical operation.

Turing's other concept addressed Lovelace's contention that machines would never think. He called it "Lady Lovelace's Objection."[ii] He asked, How would we know that? How could we test whether a machine could really think?

His answer got named the Turing Test. Put a machine and a person behind a curtain and feed them both questions, he suggested. If you cannot tell which is which, then it makes no sense to deny that the machine is thinking. This was in 1950, and he predicted that in the subsequent few decades machines would be built that would pass the Turing Test.

From Lovelace and Turing we can define two schools of thought about the relationship between humans and machines.

The Turing approach is that the ultimate goal of powerful computing is artificial intelligence: machines that can think on their own, that can learn and do everything that the human mind can do. Even everything a humanist can do.

The Lovelace approach is that machines will never truly think, and that humans will always provide the creativity and intentionality. The goal of this approach is a partnership between humans and machines, a symbiosis where each side does what it does best. Machines augment rather than replicate and replace human intelligence.

We humanists should root for the triumph of this human-machine partnership strategy, because it preserves the importance of the connection between the humanities and the sciences.

Let's start, however, by looking at how the pursuit of pure artificial intelligence — machines that can think without us — has fared.

Ever since Mary Shelly conceived Frankenstein during a vacation with Ada's father, Lord Byron, the prospect that a man-made contraption might have its own thoughts and intentions has been frightening. The Frankenstein motif became a staple of science fiction. A vivid example was Stanley Kubrick's 1968 movie, *2001: A Space Odyssey*, featuring the frighteningly intelligent and intentional computer, Hal.

Artificial intelligence enthusiasts have long been promising, or threatening, that machines like Hal with minds of their own would soon emerge and prove Ada wrong. Such was the premise at the 1956 conference at Dartmouth, organized by John McCarthy and Marvin Minsky, where the field of "artificial intelligence" was launched. The conferees concluded that a breakthrough was about twenty years away. It wasn't. Decade after decade, new waves of experts have claimed that artificial intelligence was on the visible horizon, perhaps only twenty years away. Yet true artificial intelligence has remained a mirage, always about twenty years away.

John von Neumann, the breathtakingly brilliant Hungarian-born humanist and scientist who helped devise the architecture of modern digital computers, began working on the challenge of artificial intelligence shortly before he died in 1957. He realized that the architecture of computers was fundamentally different from that of the human brain. Computers were digital and binary — they dealt in absolutely precise units — whereas the brain is partly an analog system, which deals with a continuum of possibilities. In other words, a human's mental process includes many signal pulses and analog waves from different nerves that flow together to produce not just binary yes-no data but also answers such as “maybe” and “probably” and infinite other nuances, including occasional bafflement. Von Neumann suggested that the future of intelligent computing might require abandoning the purely digital approach and creating “mixed procedures” that include a combination of digital and analog methods.[iii]

In 1958, Cornell professor Frank Rosenblatt published a mathematical approach for creating an artificial neural network like that of the human brain, which he called a “Perceptron.” Using weighted statistical inputs, it could, in theory, process visual data. When the Navy, which was funding the work, unveiled the system, it drew the type of press hype that has accompanied many subsequent artificial intelligence claims. “The Navy revealed the embryo of an electronic computer today that it expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence,” the *New York Times* wrote. The *New Yorker* was equally enthusiastic. “The Perceptron, … as its name implies, is capable of what amounts to original thought,” it reported.[iv]

That was almost sixty years ago. The Perceptron still does not exist. However, almost every year since then there have been breathless reports about some “about-to-be marvel” that would surpass the human brain, many of them using almost the exact same phrases as the 1958 stories about the Perceptron.

Discussion about artificial intelligence flared up a bit after IBM’s Deep Blue, a chess-playing machine, beat world champion Garry Kasparov in 1997, and then Watson, its natural-language question-answering cousin, won at *Jeopardy!* against champions Brad Rutter and Ken Jennings in 2011. But these were not true breakthroughs of artificial intelligence. Deep Blue won its chess match by brute force; it could evaluate 200 million positions per second and match them against 700,000 past grandmaster games. Deep Blue’s calculations were fundamentally different, most of us would agree, from what we mean by “real” thinking. “Deep Blue was only intelligent the way your programmable alarm clock is intelligent,” Kasparov said. “Not that losing to a $10 million alarm clock made me feel any better.”[v]

Likewise, Watson won at *Jeopardy!* by using megadoses of computing power: It had 200 million pages of information in its four terabytes of storage, of which the entire Wikipedia accounted for merely 0.2% of that data. It could search the equivalent of a million books per second. It was also rather good at processing colloquial English. Still, no one who watched would bet on it passing the Turing Test. For example, one question was about the “anatomical oddity” of former Olympic gymnast George Eyser. Watson answered, “What is a leg?” The correct answer was that Eyser was missing a leg. The problem was understanding an “oddity,” David Ferrucci, who ran the Watson project at IBM explained. “The computer wouldn't know that a missing leg is odder than anything else.”[vi]

Here's the paradox: Computers can do some of the toughest tasks in the world (assessing billions of possible chess positions, finding correlations in hundreds of Wikipedia-sized information repositories), but they cannot perform some of the tasks that seem most simple to us mere humans. Ask Google a hard question like, "What is the depth of the Red Sea?" and it will instantly respond 7,254 feet, something even your smartest friends don't know. Ask it an easy one like, "Can an earthworm play basketball?" and it will have no clue, even though a toddler could tell you, after a bit of giggling.[vii]

At Applied Minds near Los Angeles, you can get an exciting look at how a robot is being programmed to maneuver, but it soon becomes apparent that it still has trouble navigating across an unfamiliar room, picking up a crayon, or writing its name. A visit to Nuance Communication near Boston shows the wondrous advances in speech recognition technologies that underpin Siri and other systems, but it's also apparent to anyone using Siri that you still can't have a truly meaningful conversation with a computer, except in a fantasy movie. A visit to the New York City police command system in Manhattan reveals how computers scan thousands of feeds from surveillance cameras as part of a "Domain Awareness System," but the system still cannot reliably identify your mother's face in a crowd.

There is one thing that all of these tasks have in common: even a four-year-old child can do them.

Perhaps in a few more decades there will be machines that think like, or appear to think like, humans. "We are continually looking at the list of things machines cannot do — play chess, drive a car, translate language — and then checking them off the list when machines become capable of these things," said Tim Berners-Lee, who invented the World Wide Web. "Someday we will get to the end of the list."[viii]

Someday we may even reach the "singularity," a term that John von Neumann coined and the science fiction writer Vernor Vinge popularized, which is sometimes used to describe the moment when computers are not only smarter than humans but can also design themselves to be even super smarter, and will thus no longer need us mere mortals. Vinge says this will occur by 2030.[ix]

On the other hand, this type of artificial intelligence may take a few more generations or even centuries. We can leave that debate to the futurists. Indeed, depending on your definition of consciousness, it may never happen. We can leave that debate to the philosophers and theologians.

Endnotes

[i] Ada, Countess of Lovelace, "Notes on Sketch of The Analytical Engine," October, 1842.
[ii] Alan Turing, "Computing Machinery and Intelligence," *Mind*, October 1950.
[iii] John von Neumann, *The Computer and the Brain* (Yale, 1958), 80.
[iv] Gary Marcus, "Hyping Artificial Intelligence, Yet Again," *New Yorker*, Jan. 1, 2014, citing: "New Navy Device Learns by Doing", (UPI wire story) *New York Times*, July 8, 1958; "Rival", *New Yorker*, Dec. 6, 1958.
[v] Garry Kasparov, "The Chess Master and the Computer," *The New York Review of Books*, Feb. 11, 2010; Clive Thompson, *Smarter Than You Think* (Penguin, 2013), 3.
[vi] Watson on Jeopardy, on IBM's Smarter Planet Website, Feb. 14, 2011.
[vii] Gary Marcus, "Why Can't My Computer Understand Me," *New Yorker*, Aug. 16, 2013, coined the example I adapted.
[viii] Author's interview with Tim Berners-Lee.
[ix] Vernor Vinge, "The Coming Technological Singularity," *Whole Earth Review*, Winter 1993.

[*Note: This essay version has been partitioned and subtitled by the editors in order to facilitate student interaction. It may not conform in all details to the spoken lecture or transcript.*]