Is passing a Turing Test a true measure of artificial intelligence?

11 June 2014, by Kevin Korb

What does it take for a computer to show artificial intelligence? Credit: Flickr/Nebraska Oddfish, CC BY-NC-SA

The Turing Test has been passed, the headlines report this week, after a computer program mimicked a 13-year-old Ukrainian boy called Eugene Goostman, fooling 33% of its interrogators into believing it was human after five minutes of questioning.

But this isn't the first time the test has been "passed", and there remain questions of its adequacy as a test of artificial intelligence.

The Turing Test came about in 1950 when British mathematician and codebreaker Alan Turing wrote a provocative and persuasive article, Computing Machinery and Intelligence, advocating for the possibility of artificial intelligence (AI).

Having spent the prior decade arguing with psychologists and philosophers about that possibility, as well as helping to crack the Nazi Enigma machine at Bletchley Park in England, he became frustrated with the prospect of actually *defining* intelligence.

So he proposed instead a behavioural criterion along the following lines:

If a machine could fool humans into thinking it's a human, then it must be at least as intelligent as a normal human.

Turing, being an intelligent man, was more cautious than many AI researchers who came after, including Herbert "10 Years" Simon, predicting in 1957 success in 1967.

Turing predicted that by the year 2000 a program would be made which would fool the "average interrogator" 30% of the time after five minutes of questioning.

These were not meant as defining conditions on his test, but merely as an expression of caution. It turns out that he too was insufficiently cautious, getting it right at the end of his article: "We can see plenty there that needs to be done."

What first passed the Turing Test?

Arguably the first was ELIZA, a program written by the American computer scientist Joseph Weizenbaum.

His secretary was fooled into thinking she was communicating with him remotely, as described in his 1976 book Computer Power and Human Reason.

Weizenbaum left his program running at a terminal, which she assumed was connected to Weizenbaum himself, remotely.

In conversation with Eugene Goostman. Credit: Princeton AI

Subsequent programs which fooled humans include "PARRY", which "pretended" to be a paranoid schizophrenic.

In general, it has not gone unnoticed that programs that mimic those with a limited behavioural repertoire, limited knowledge or understanding, or engage those who are predisposed to accept their authenticity are more likely to fool their audience.

But none of this engages the real issues of artificial intelligence today:

1. what criterion would establish something close to human-level intelligence?
2. when will we achieve it?
3. what are the consequences?

The criterion

The Turing Test, even as envisaged by Turing, let alone as manipulated by publicity seekers, has limitations.

As US philosopher John Searle and cognitive scientist Stevan Harnad have already pointed out, anything like human intelligence must be able to engage with the real world ("symbol grounding"), and the Turing Test doesn't test for that.

My view is that they are right, but that passing a genuine Turing Test would nevertheless be a major achievement, sufficient to launch the Technological Singularity – the point when intelligence takes off exponentially in robots.

The timeframe

We will achieve AI in 1967 predicted Herbert Simon, or 2000 suggested Alan Turing, or 2014 with the Eugene Goostman program on the weekend, or much later. All the dates before 2029 are in my view just silly.

Google's director of engineering Ray Kurzweil at least has a real argument for 2029, based on Moore's Law-type progress in technological improvement.

However, his arguments don't really work for software. Progress in improving our ability to design, generate and test software has been comparatively painfully slow.

As IBM's Fred Brooks famously wrote in 1986, there is "No Silver Bullet" for software—nor is there now. Modelling (or emulating) the human brain, with something like 10^{14} synapses, would be a software project many orders of magnitude larger than the largest software project ever done.

I consider the prospect of organising and completing such a project by 2029 to be remote, since this appears to be a project of greater complexity than any human project ever undertaken. An estimate of 500 years to complete it seems to me far more reasonable.

Of course, I might have said the same thing about some hypothetical "Internet" were I writing in Turing's time. In general, scheduling (predicting) software is one of the mysteries no one seems to have mastered.

The consequences

The consequences of passing the true Turing Test and achieving a genuine Artificial Intelligence will be massive.

As Irving John Good, a coworker of Turing at Bletchley Park, pointed out in 1965, a general AI could be put to the task of improving itself, leading to rapidly increasing improvements recursively, so that "the first ultraintelligent machine is the last invention that [humans] need ever make."

This is the key to launching the Technological

Singularity, the stuff of Hollywood nightmares and Futurist dreams.

While I am sceptical of any near-term Singularity, I fully agree with Australian philosopher David Chalmers who argues that the consequences are sufficiently large that we should even now concern ourselves with the ethics of the Singularity.

Famously, science fiction author Isaac Asimov (implicitly) advocated enslaving AIs with his three laws of robotics, binding them to do our bidding. I consider enslaving intelligences far greater than our own of dubious merit, ethically or practically.

More promising would be building a genuine ethics into our AI, so that they would be unlikely to fulfill Hollywood's fantasies.

This story is published courtesy of The Conversation *(under Creative Commons-Attribution/No derivatives).*

Source: The Conversation