Tests That Show Machines Closing in on Human Abilities

By Colin Barras, New Scientist

Jan. 22, 2009 — It may have been dreamt up in 1950, but the Turing test - a simple way to tell if a machine can think - still holds powerful sway over many researchers striving to produce a machine at least in some respects equal with a human. Nowadays, although UK mathematician Alan Turing's test is still relevant, and unbeaten, new forms of it have evolved. In this online special, New Scientist discovers the different ways in which machines can be tested for human-like abilities - and how close they have come to passing as one of us.

Video: As new ways to test how well machines can match aspects of human intelligence are dreamt up, they are getting closer to beating them.

I chat, therefore I think

Turing's biggest insight was that it is impossible to know for sure if a machine - or indeed another person - is actually thinking. So he rephrased the question to one that is much easier to answer: Can a machine act like it is thinking?

His test has a human judge engage in two, separate conversations using only text - one with a human and one with a machine. If, after a few minutes of conversation, the judge can't distinguish the two, then the machine is deemed to have passed the test.

A major turning point in the field was Joseph Weizenbaum's chatbot ELIZA, developed in 1966. It achieved impressively "intelligent" results simply by rephrasing what a human said to it:

Human: My boyfriend made me come here.

ELIZA: Your boyfriend made you come here.

Human: He says I'm depressed much of the time.

ELIZA: I am sorry to hear you are depressed.

Human: It's true. I am unhappy.

ELIZA: Do you think coming here will help you not to be unhappy?

That technique was modelled on the way psychotherapists attempt to draw out a patient.

"If, for example, one were to tell a psychiatrist 'I went for a long boat ride' and he responded 'Tell me about boats', one would not assume that he knew nothing about boats, but that he had some purpose in so directing the subsequent conversation," wrote Weizenbaum in his 1966 paper.

He realised that relatively simple strategies can convince a human they are talking with another intelligence - some people found it very difficult to believe that ELIZA wasn't a human.

The 2008 winner of annual Turing test contest the Loebner Prize also won using a brilliantly simple strategy that Turing didn't foresee. Elbot convinced three of 12 humans it was just like them by acting like a human pretending to be a robot.

Escaping the Uncanny Valley

Turing lived in an age when machines were room-filling boxes, but today's robotic creations are edging towards looking passably human. In fact, they are solidly inside what is known as "uncanny valley", a phrase that describes the creepy look of objects that look close, but not close enough, to being human.

The first artificial humans to emerge from the valley are likely to be creations of computer animators; but even they all too easily get uncanny results when using the latest tools. Last summer, though, US animation company Image Metrics and Paul Debevec at the University of Southern California produced a character some claimed had truly pass a visual Turing test (see video above).

Actress Emily O'Brien's skin was scanned in high-resolution as she changed her facial expressions. The results were turned into an animated version of O'Brien's face that was detailed enough to show facial pores and fine wrinkles. It was overlaid onto video of her real body to make a short movie realistic enough to defy the uninformed viewer.

The line between reality and animation is blurring, and improved models of themuscles and tendons of the real human body will accelerate the process.

Hiroshi Ishiguro, a roboticist at Osaka University in Japan, may be the person closest to traversing uncanny valley physically. His androids include creepily accurate replicas of himself and his daughter, complete with sensitive skin made from shape-memory plastic and plaster.

His doppelganger - Geminoid - travels with him to conferences and even lectures in his place. Its looks, combined with its ability to fidget like a human, have proved enough to fool some people it is one of them.

Ishiguro asked 20 volunteers to watch a curtain rise in front of them, and note the colour of a piece of cloth behind it. To their surprise, the opening curtain also revealed one of Ishiguro's androids, sat motionless. Most of the participants - 70% - reported afterwards that the figure was artificial.

But when the humanoid reproduced the tiny "micro-movements" produced subconsciously by a seated human, and the test was repeated with more volunteers, 70% of them believed the figure was a person.

Future robots may use other strategies to look real. Seth Goldstein at Carnegie Mellon University, Pittsburgh, last year took the first speculative stepstowards creating swarms of tiny robots that cling together using electrostatic forces to take on any form - technology that has been dubbed "claytronics".

"I'll be done when we produce something that can pass a Turing test for appearance," he told New Scientist. "You won't know if you're shaking hands with me or a claytronics copy of me."

Sounds real good

François Pachet, a researcher at Sony Computer Science Laboratories, Paris, is a frustrated jazz musician - despite years of mastering scales, improvisation always eluded him.

So, in 1999, he decided to build a robot able to improvise like a pro. HisContinuator can duet with a live musician in real-time.

It "listens" to a musical phrase and then computes a complementary phrase with the same playing style. Pachet worked with professional musicians to develop and improve Continuator and by 2002 it could respond to polyphonic melodies with intricate melodies of its own.

Then in 2004, Continuator faced up to a kind of musical Turing test. Pachet asked two jazz critics to listen to jazz pianist Albert Van Veenendaalimprovising on a piano hooked up to the Continuator.

When the expert played a short refrain, Continuator responded. But the critics found it impossible to reliably distinguish who was playing.

You can see Continuator tinkling the virtual ivories in the video above, or try and tell man from machine in the audio clips here.

Decisions of life and death

The tests mentioned so far may be significant milestones, but robotic systems may soon have to make decisions over life and death. At a conference last year legal and military experts suggested entrusting a Turing test with an altogether more serious job - determining whether or not military equipment may choose for itself when and where to fire.

Debate at the Ethics of Autonomous Military Systems conference centred on whether a robot could yet commit a war crime. It was generally agreed that this could occur only when robotic systems become so autonomous that no human can foresee their actions, and therefore no human could be culpable for the robot's behaviour.

That point is still far off. And legislation may prevent any future machine with that capability from being deployed, said Chris Elliot, an engineer and barrister. "Weapons intrinsically incapable of distinguishing between civilian and military targets are illegal," he says. However, Ronald Arkin, of Georgia Tech University, argues that machines will one day be able to make that judgement as well as a human can. One that does will have passed the military Turing test.

Quite how such a military Turing test might be validated safely is a moot point. But Arkin believes that such machines could even be more ethical soldiers than humans - free as they are from emotions and prejudice over human traits such as race.

Watching the brain

The tests above have involved asking humans what they thought about robot performance. But some experts think that analysing how people think when they interact with machines is a better approach.

In 2007, Lindsay Oberman and colleagues at the University of California, San Diego, monitored the brain activity of 17 students as they watched video of a five-fingered metallic robot hand or control footage of static noise.

Only watching the hand triggered activity in the students' mirror neurons, cells which fire when a person sees someone performing a physical action they might perform themselves.

"If we want humanoid robots to teach or have other social functions, we need them to trigger mirror neurons," Oberman told New Scientist.

In 2008, Sören Krach, a psychiatrist at Aachen University in Germany, reported similar results. His experiments monitored the brain activity of 20 students playing a simple cooperation game with either a laptop, a pair of robotic arms, a child robot with a rubber head, or a human for a partner.

The more a player's partner resembled a human, the greater the neural activity observed. Players also found the game more interesting when playing with the more human partners, and reported them more intelligent, despite all the artificial players having identical underlying software (PLoS One, DOI: 10.1371/journal.pone.0002597).

Game over

Few humans would begrudge a machine due praise for passing any of these tests. But acknowledging the consequences of passing a Turing test are a little harder to stomach.

Cognitive scientist Stevan Harnad at the University of Southampton, UK, points out that we should be open minded enough to recognise the truth of such a result, and not dismiss it as a "trick":

"AI is not a party game, the game was just a metaphor," he says. "If a machine can prove indistinguishable from a human, we should award it the respect we would to a human - we should accept that it has a mind."