

By Invitation

# We, the Robots?

As computers surpass human intelligence and take on greater responsibilities, should they be given rights as well?



Simon Chesterman

For *The Straits Times*

Earlier this month, the Bank of England announced that Alan Turing would be featured on the new £50 (S\$84.60) note. The last to switch from paper to polymer, it is due to enter circulation in late 2021.

Turing was a hero for his codebreaking during World War II. He also helped to establish the discipline of computer science, laying the foundations for what we now call artificial intelligence (AI).

Among his contributions is a test for when true "intelligence" has actually been achieved. He modelled it on a parlour game that was popular in 1950. A man and a woman sit in a separate room and provide written answers to questions; the other participants have to guess who provided which answer. Turing posited that a similar "imitation game" could be played with a computer. When a machine could fool people into believing that it was human, we might properly say that it was intelligent.

## I THINK, THEREFORE...

Parlour games aside, why should we care about whether a computer is "intelligent"?

As AI systems become more sophisticated, there are at least two reasons why we might start thinking of them as intelligent – and, indeed, as persons in their own right.

The first is so that we have someone to blame when things go wrong.

Ten years from now, if an autonomous vehicle crashes into you, it may be pointless to fault the human "driver", as the car may not even have a steering wheel. Law reform bodies are already suggesting the concept of an automated driving system entity, or ADSE, to govern liability questions.

Similarly, if an autonomous agent – say, a robot security guard – breaks the law, it may be impossible to tie the criminal conduct to an identifiable person or corporation. Some experts have proposed that such machines could be "punished" by being reprogrammed or, in extreme cases, destroyed.

The second reason is so that we have someone to reward when things go right.

Not long ago, a novella written by AI was shortlisted in a Japanese literary contest. Computers are increasingly engaging in quasi-creative activities. Who



should own such creations?

Consider the world's most famous selfie – of a crested black macaque. Mr David Slater went to Indonesia to photograph the endangered monkeys, which were too nervous to let him take close-ups. So he set up a camera that enabled them to snap their own photos.

But who owned the images: Mr Slater? The monkeys? No one?

Mr Slater did eventually win, reflecting existing law. But, as computers generate more content independently of their human programmers, it is going to be harder and harder for us to take credit. Instead of teaching a monkey how to press a button, it will be more like a teacher trying to take credit for the work of his or her student.

## WE HOLD THESE TRUTHS TO BE SELF-EVIDENT...

If one believes that some kind of personhood is appropriate, how should we decide which machines get it?

The Turing Test offers one approach. If people can't tell the difference between a machine and a human, maybe that's when the machine deserves to be treated like one.

The earliest successes in the Turing Test came in the 1960s with

programs such as Eliza.

Users were told that Eliza was a psychotherapist who communicated through words typed into a computer. In fact, "she" was a program using a simple list-processing language. If the user typed in a recognised phrase, it might be reframed as a question. So after entering "I'm depressed", Eliza might reply "Why do you say that you are depressed?" If it didn't recognise the phrase, it would offer something generic, such as "Can you elaborate on that?"

Even when users were told how it worked, some insisted that Eliza had "understood" them. Nevertheless, no one today would argue that Eliza should be treated as a "person".

So how should we think about the personhood of our metal and silicon counterparts?

A major hurdle is that the more we learn about machine intelligence, the less it seems we understand our own intelligence. That is all the more true of what we call "consciousness".

But as a legal category, we do have ways of approaching the question of personality.

There are essentially two types of person recognised by law. Natural persons – humans like you and me – are recognised because of the simple fact of being human.



The famous "monkey selfie" raises questions about the ownership of content generated independently of humans. PHOTO: WIKIMEDIA COMMONS

Juridical persons, by contrast, are non-human entities – companies, for example – that are granted certain duties and rights by law.

It might seem self-evident that a computer can't be a natural person. Yet, for centuries, slaves and women weren't recognised as full natural persons either.

If we take the Turing Test to its natural, Blade Runner conclusion, it is possible that AI systems truly indistinguishable from humans

might claim the same status.

For the time being, however, we are likely to stick with the possibility of juridical personhood.

Unlike natural persons, the law does not recognise juridical persons for their inherent qualities but for their instrumental value. Limited liability corporations were created to encourage entrepreneurship. If sued for wrongdoing, the corporation itself is liable, rather than its investors.

Such arguments might apply to certain forms of AI. Given the diversity of AI systems, however, it is simplistic to lump them all into a single category of "person".

Implicit in many of those arguments is an assumption that AI research is on a path towards some kind of AI consciousness and a claim to natural personhood. That will probably remain science fiction – or, at least, it is not a sensible basis for regulation today.

The better solution is to rely on existing categories, with responsibility for wrongdoing tied to users, owners or manufacturers, rather than the AI systems themselves. Driverless cars are following this path, for example, with a likely shift from insuring drivers to insuring vehicles.

As for ownership of intellectual property created by AI, the moral and economic reasons for

protecting human creations don't seem to apply. Would it be unjust, or would it stop AI systems doing creative work if they "knew" that their output could be used by anyone?

## MACHINES LIKE US

In the 1990s, a prize was established to encourage more serious attempts at the Turing Test.

One of the first winners succeeded in part by tricking people – the program made spelling mistakes that testers assumed must have been the result of human fallibility.

Though the Turing Test remains a cultural touchstone, it is far from the best measure of AI research today. As a leading textbook notes, the quest for flight succeeded when the Wright brothers stopped trying to imitate birds and started learning about aerodynamics.

Aeronautical engineers today don't define the goal of their field as making machines that fly so exactly like pigeons that they can fool other pigeons.

Turing himself never lived to see a computer even attempt his test. Prosecuted for homosexual acts in 1952, he chose chemical castration as an alternative to prison. He died two years later at the age of 41, apparently after committing suicide by eating a cyanide-laced apple.

The announcement that Turing will grace the new £50 note follows an official pardon signed by the Queen in 2013.

Yet the more fitting tribute may be Ian McEwan's most recent novel, *Machines Like Me*, which imagines an alternative timeline in which Turing lived and was rewarded with the career and the knighthood he deserved.

The novel takes seriously the prospect of true artificial intelligence in the form of a brooding synthetic Adam, who expresses his love for the human Miranda by writing thousands upon thousands of haikus.

Ultimately, however, consciousness is a burden for the machines – struggling to find their place in the world, and so binary that they are unable to reconcile human virtues and human vices.

It also offers Turing a chance to rethink his test. "In those days," the fictional Turing says at age 70, referring to his younger self, "I had a highly mechanistic view of what a person was. The body was a machine, an extraordinary one, and the mind I thought of mostly in terms of intelligence, which was best modelled by reference to chess or maths."

The reality, of course, is that chess is not a representation of life. Life is an open system; it is messy. It is also unpredictable.

In the novel, the first priority of the AI robots is to disable the kill switch that might shut them down. Yet most of them ultimately destroy themselves – as the real Turing did – unable to reconcile their innate nature with the injustices of the world around them.

Before asking whether we could create such thinking machines, McEwan reminds us, we might want to pause and ask whether we should.

stopinion@sph.com.sg

• Simon Chesterman is Dean and Professor of the National University of Singapore's Faculty of Law.