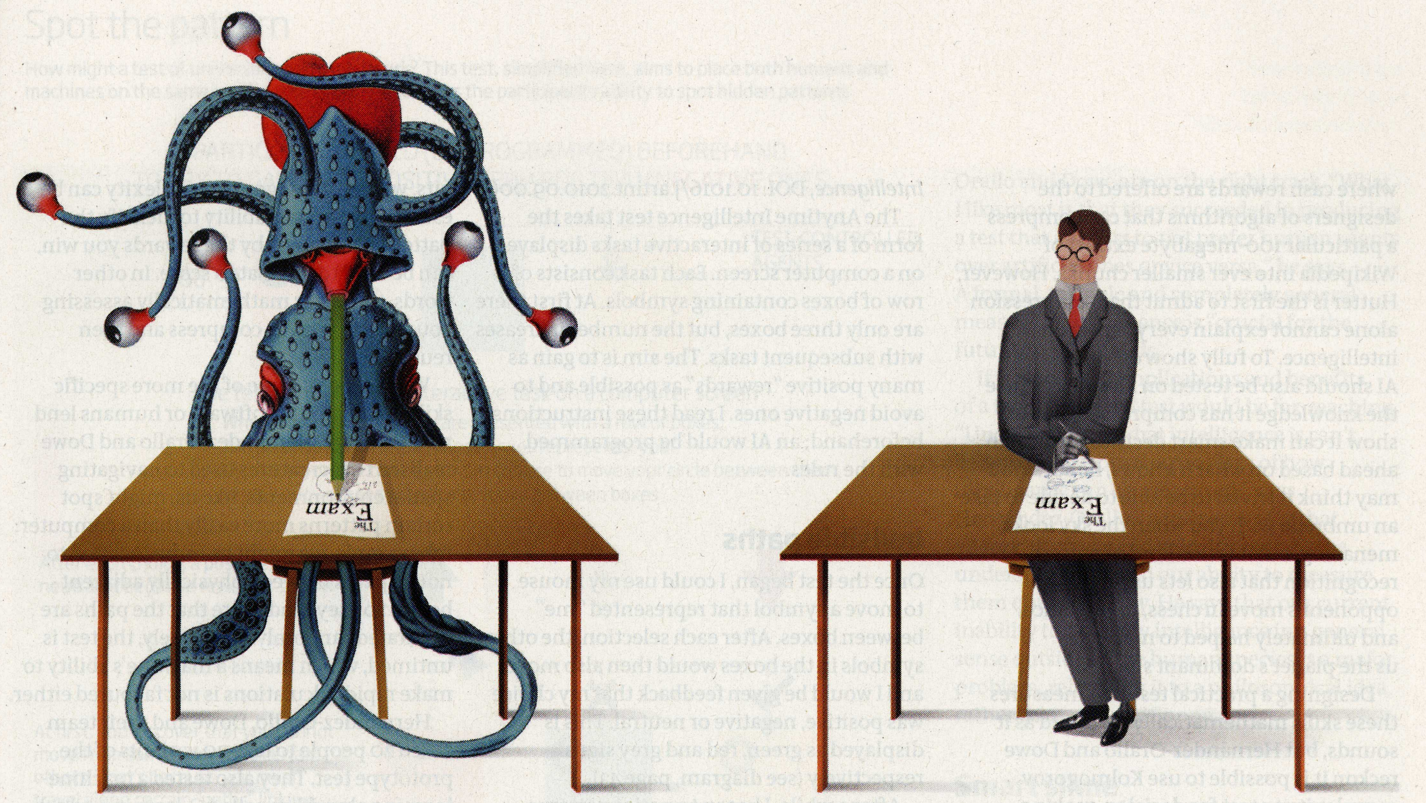

definition of intelligence, and often assume knowledge and skills that are unique to us.

For AI, the best-known performance test is Turing's. However, finding a program that can imitate the abilities of the human mind has proved to be a big challenge. Since 1990, the Loebner prize competition, based on the Turing test, has sparked the creation of a multitude of "chatbots" with fairly impressive social skills, yet most AI researchers don't think these are truly smart. "The Turing test leads to interesting philosophical arguments about intelligence in general and how can we measure it – but it has never been taken seriously as the ultimate goal for AI," says Marcus Hutter at the Australian National University in Canberra.

Instead, myriad specific tests measure "narrow" types of AI – for example, the ability to play chess. IBM's Deep Blue beat Garry Kasparov in 1997, and yet it would be utterly useless at adapting itself to complete a crossword or even to figure out the best way to fold your clothes.

One attempt to encourage the development of AIs with broader intelligence is the General Game Playing competition, held annually at the meeting of the American Association for Artificial Intelligence. Bots are served up a combination of games, from noughts and crosses (tic-tac-toe) to draughts (checkers). They must then devise their own game plans using nothing but a list of rules for those games given to them beforehand.

However, the contest is still asking machines to play at being human. Could

"If we encounter aliens, an intelligence test based on mathematics would tell us what we're dealing with"

there be another, independent, benchmark for their intelligence? If so, we could compare machines with each other much more accurately – as well as with ourselves.

Hernández-Orallo decided to devise such a test, along with David Dowe, who specialises in information theory and statistics at Monash University in Melbourne, Australia. For inspiration, they turned to a mathematical definition of intelligence with its roots in the 1960s.

Back then, AI pioneer Ray Solomonoff related intelligence to the ability to summarise or "compress" information by detecting patterns. This skill allows for better problem-solving than using mere trial and error. For example, faced with the sequences 10101010101010 or 1234567, a machine or person that realises these can be summarised as "repeat '10' seven times" or "count to 7" is rated as more intelligent than one that doesn't. Compression also leads to the ability to predict: a machine that can spot the pattern can use that information to name subsequent digits. This is related to predictive learning – essentially the ability to learn by spotting, generalising and reusing patterns.

It's nothing new that finding patterns is related to intelligence. But Solomonoff's contribution was to mathematically quantify the process of predictive learning, using a concept now known as Kolmogorov complexity. Information that can be easily compressed – such as the sequences above – has low Kolmogorov complexity, whereas a truly random sequence, which cannot be compressed at all, has high Kolmogorov complexity. Despite their implications for AI, Solomonoff's ideas were largely ignored until the late 1990s, when Dowe and later Hernández-Orallo began to explore the connection between compression and intelligence.

In an effort to spur AI researchers further in this direction, in 2006 Hutter launched the Human Knowledge Compression prize, commonly known as the Hutter prize, ›