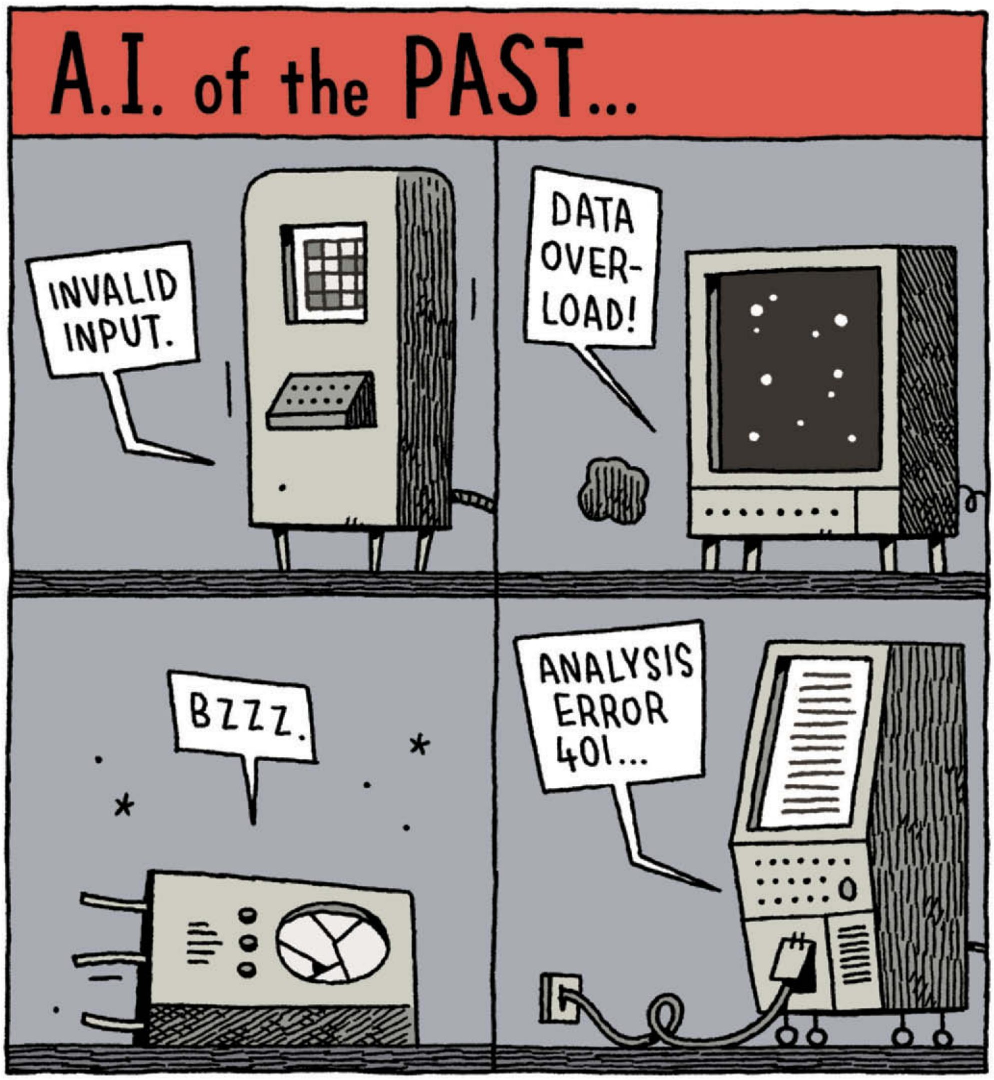


At last, artificial intelligences are thinking along human lines. Anil Ananthaswamy reports

I, algorithm

GIVEN the choice between a flesh-and-blood doctor and an artificial intelligence system for diagnosing diseases, Pedro Domingos is willing to stake his life on AI. "I'd trust the machine more than I'd trust the doctor," says Domingos, a computer scientist at the University of Washington, Seattle. Considering the bad rap AI usually receives – overhyped, underwhelming – such strong statements in its support are rare indeed.

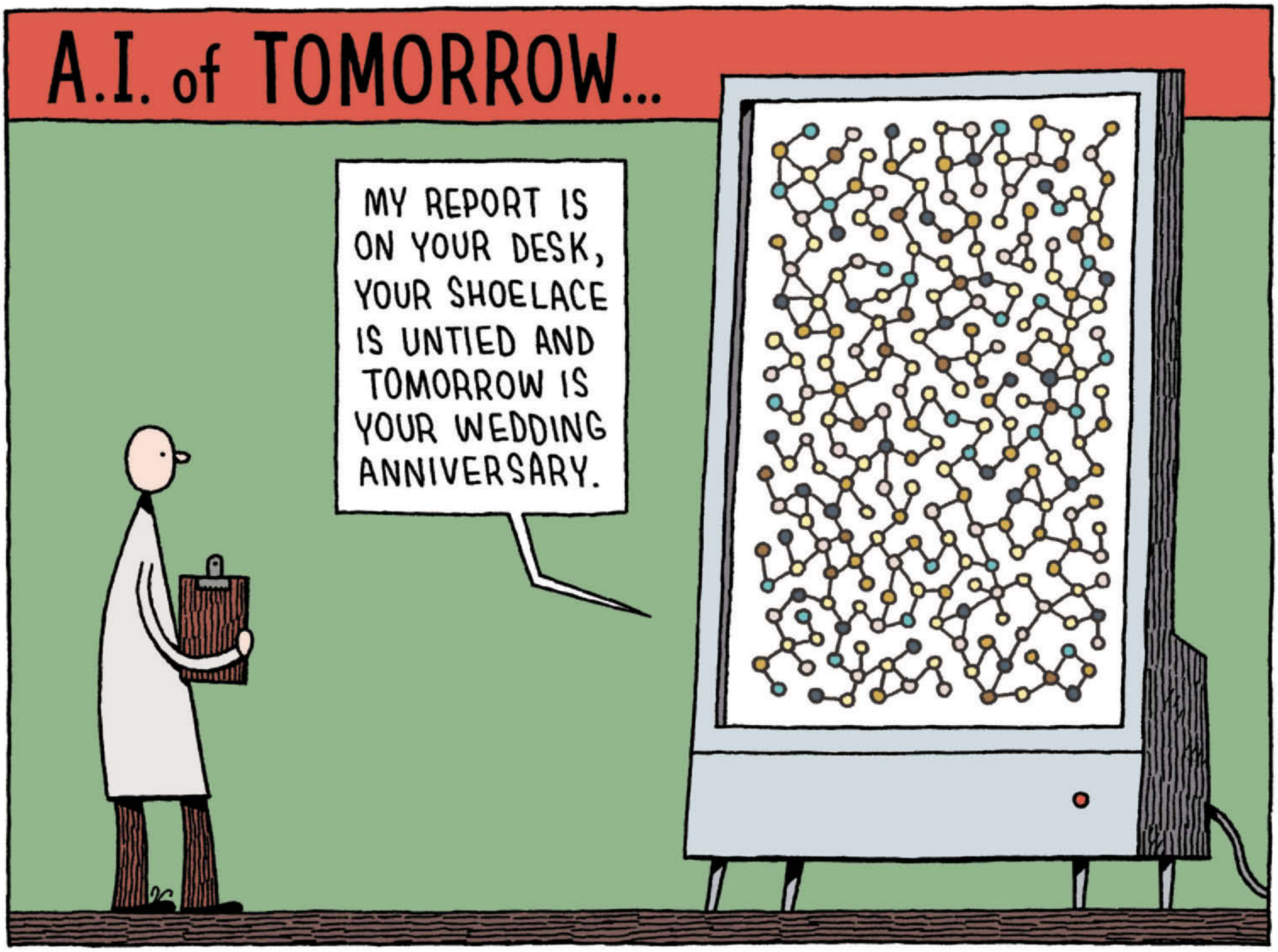
Back in the 1960s, AI systems started to show great promise for replicating key aspects of the human mind. Scientists began by using mathematical logic to both represent knowledge about the real world and to reason about it, but it soon turned out to be an AI straightjacket. While logic was capable of being productive in ways similar to the human mind, it was inherently unsuited for dealing with uncertainty.

Yet after spending so long shrouded in a self-inflicted winter of discontent, the much-maligned field of AI is in bloom again. And Domingos is not the only one with fresh confidence in it. Researchers hoping to detect illness in babies, translate spoken words into text and even sniff out rogue nuclear explosions are proving that sophisticated computer systems can exhibit the nascent abilities which sparked interest in AI in the first place: the ability to reason like humans, even in a noisy and chaotic world.

Lying close to the heart of AI's revival is a technique called probabilistic programming, which combines the logical underpinnings of the old AI with the power of statistics and probability. "It's a natural unification of two of the most powerful theories that have been developed to understand the world and reason about it," says Stuart Russell, a pioneer of modern AI at the University of California, Berkeley. This powerful combination is finally starting to disperse the fog of the long AI winter. "It's definitely spring," says cognitive scientist Josh Tenenbaum at the Massachusetts Institute of Technology.

The term "artificial intelligence" was coined in 1956 by John McCarthy of MIT. At the time, he advocated the use of logic for developing computer systems capable of reasoning. This approach matured with the use of so-called first-order logic, in which knowledge about the real world is modelled using formal mathematical symbols and notations. It was designed for a world of objects and relations between objects, and it could be used to reason about the world and arrive at useful conclusions. For example, if person X has disease Y, which is highly infectious, and X

TOM GOULD



came in close contact with person Z, using logic one can infer that Z has disease Y.

However, the biggest triumph of first-order logic was that it allowed models of increasing complexity to be built from the smallest of building blocks. For instance, the scenario above could easily be extended to model the epidemiology of deadly infectious diseases and draw conclusions about their progression. Logic's ability to compose ever-larger concepts from humble ones even suggested that something analogous might be going on in the human mind.

That was the good news. "The sad part was that, ultimately, it didn't live up to expectations," says Noah Goodman, cognitive scientist at Stanford University in California. That's because using logic to represent knowledge, and reason about it, requires us to be precise in our know-how of the real world. There's no place for ambiguity. Something is either true or false, there is no maybe. The real world, unfortunately, is full of uncertainty, noise and exceptions to almost every general rule. AI systems built using first-order logic simply failed to deal with it. Say you want to tell whether person Z has disease Y. The rule has to be unambiguous: if Z came into contact with X, then Z has disease Y. First-order logic cannot handle a scenario in which Z may or may not have been infected.

There was another serious problem: it didn't work backwards. For example, if you knew that Z has disease Y, it was not possible to infer with absolute certainty that Z caught it from X. This typifies the problems faced in medical diagnosis systems. Logical rules can link diseases to symptoms, but a doctor faced with symptoms has to infer backwards to the cause. "That requires turning around the logic formula, and deductive logic is not a very good way to do that," says Tenenbaum.

These problems meant that by the mid-1980s, the AI winter had set in. In popular perception, AI was going nowhere. Yet Goodman believes that, secretly, people didn't give up on it. "It went underground," he says.

The first glimmer of spring came with the arrival of neural networks in the late 1980s. The idea was stunning in its simplicity. Developments in neuroscience had led to simple models of neurons. Coupled with advances in algorithms, this let researchers build artificial neural networks (ANNs) that could learn, ostensibly like a real brain. Invigorated computer scientists began to dream of ANNs with billions or trillions of neurons. ›

Yet it soon became clear that our models of neurons were too simplistic and researchers couldn't tell which of the neuron's properties were important, let alone model them.

Neural networks, however, helped lay some of the foundations for a new AI. Some researchers working on ANNs eventually realised that these networks could be thought of as representing the world in terms of statistics and probability. Rather than talking about synapses and spikes, they spoke of parameterisation and random variables. "It now sounded like a big probabilistic model instead of a big brain," says Tenenbaum.

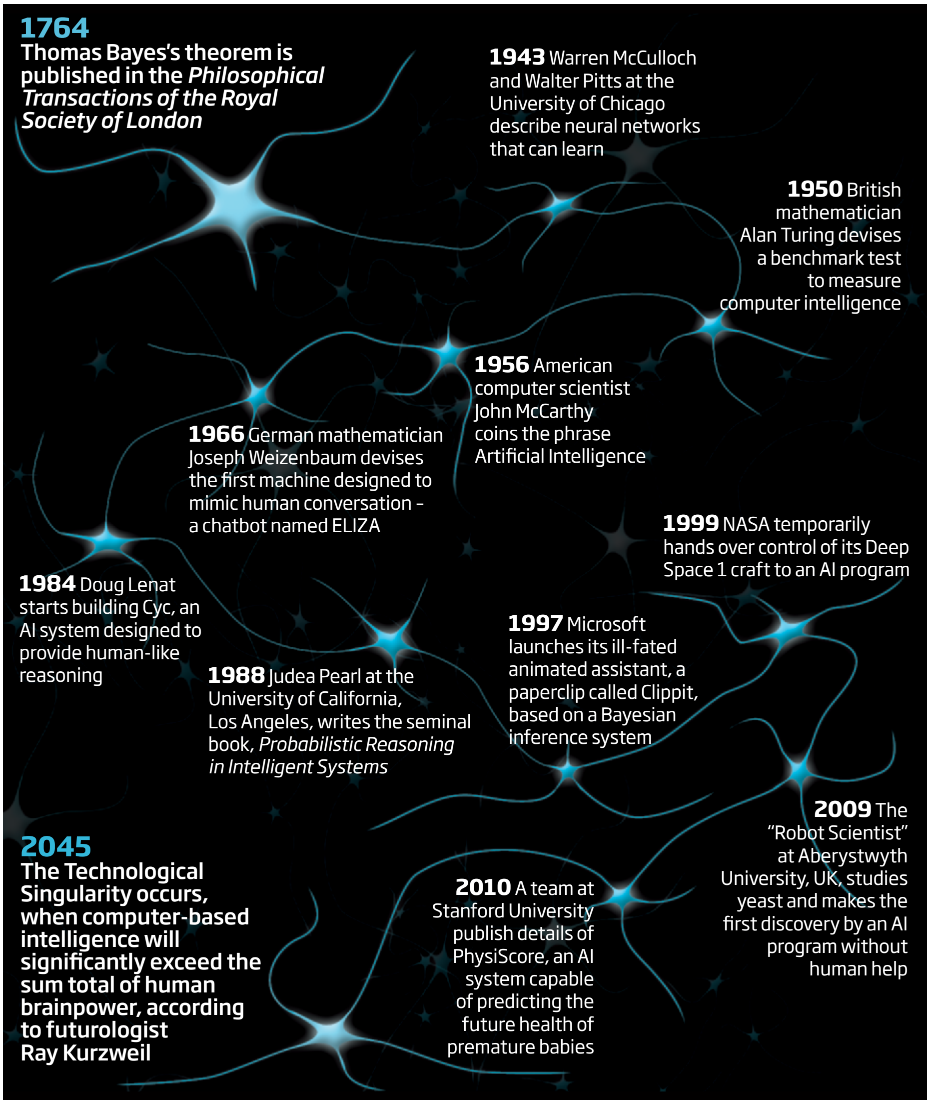
Then, in 1988, Judea Pearl at the University of California, Los Angeles, wrote a landmark book called *Probabilistic Reasoning in Intelligent Systems*, which detailed an entirely new approach to AI. Behind it was a theorem developed by Thomas Bayes, an 18th-century English mathematician and clergyman, which links the conditional probability of an event P occurring given that Q has occurred to the conditional probability of Q given P. Here was a way to go back-and-forth between cause and effect. "If you can describe your knowledge in that way for all the different things you are interested in, then the mathematics of Bayesian inference tells you how to observe the effects, and work backwards to the probability of the different causes," says Tenenbaum.

The key is a Bayesian network, a model made of various random variables, each with a probability distribution that depends on every other variable. Tweak the value of one, and you alter the probability distribution of all the others. Given the value of one or more variables, the Bayesian network allows you to infer the probability distribution of other variables – in other words, their likely values. Say these variables represent symptoms, diseases and test results. Given test results (a viral infection) and symptoms (fever and cough), one can assign probabilities to the likely underlying cause (flu, very likely; pneumonia, unlikely).

By the mid-1990s, researchers including Russell began to develop algorithms for Bayesian networks that could utilise and learn from existing data. In much the same way as human learning builds strongly on prior understanding, these new algorithms could learn much more complex and accurate models from much less data. This was a huge step up from ANNs, which did not allow for prior knowledge; they could only learn from scratch for each new problem.

The long road to artificial intelligence

Will a new take on Thomas Bayes's 250-year-old theorem clear the path to AI that can reason like a human?



Nuke hunting

The pieces were falling into place to create an artificial intelligence for the real world. The parameters of a Bayesian network are probability distributions, and the more knowledge one has about the world, the more useful these distributions become. But unlike systems built with first-order logic, things don't come crashing down in the face of incomplete knowledge.

Logic, however, was not going away. It turns out that Bayesian networks aren't enough by themselves because they don't allow you to build arbitrarily complex constructions out of simple pieces. Instead it is the synthesis of logic programming and Bayesian networks into the field of probabilistic programming that is creating a buzz.

At the forefront of this new AI are a handful of computer languages that incorporate both elements, all still research tools. There's Church, developed by Goodman, Tenenbaum and colleagues, and named after Alonzo Church who pioneered a form of logic for computer programming. Domingos's team has developed Markov Logic Network, combining Markov networks – similar to Bayesian networks – with logic. Russell and his colleagues have the straightforwardly named Bayesian Logic (BLOG).

Russell demonstrated the expressive power of such languages at a recent meeting of the UN's Comprehensive Test Ban Treaty Organization (CTBTO) in Vienna, Austria. The CTBTO invited Russell on a hunch that the new AI techniques might help with the problem of detecting nuclear explosions. After a morning listening to presenters speak about the challenge of detecting the seismic signatures of far-off nuclear explosions amidst the background of earthquakes, the vagaries of signal propagation through the Earth, and noisy detectors at seismic stations worldwide, Russell sat down to model the problem using probabilistic programming (*Advances in Neural Information Processing Systems*, vol 23, MIT Press). "And in the lunch hour I was able to write a complete model of the whole thing,"

Intelligent machines must distinguish bombs from earthquakes

JUNG YEON-JE/AFP/GETTY IMAGESE

says Russell. It was half a page long.

Prior knowledge can be incorporated into this kind of model, such as the probability of an earthquake occurring in Sumatra, Indonesia, versus Birmingham, UK. The CTBTO also requires that any system assumes that a nuclear detonation occurs with equal probability anywhere on Earth. Then there is real data – signals received at CTBTO's monitoring stations. The job of the AI system is to take all of this data and infer the most likely explanation for each set of signals.

Therein lies the challenge. Languages like BLOG are equipped with so-called generic inference engines. Given a model of some real-world problem, with a host of variables and probability distributions, the inference engine has to calculate the likelihood of, say, a nuclear explosion in the Middle East, given prior probabilities of expected events and new seismic data. But change the variables to represent symptoms and disease and it then must be capable of medical diagnosis. In other words its algorithms must be very general. That means they will be extremely inefficient.

The result is that these algorithms have to be customised for each new challenge. But you can't hire a PhD student to improve the algorithm every time a new problem comes along, says Russell. "That's not how your brain works; your brain just gets on with it."

This is what gives Russell, Tenenbaum and others pause, as they contemplate the future of AI. "I want people to be excited but not feel as if we are selling snake oil," says Russell. Tenenbaum agrees. Even as a scientist on the right side of 40, he thinks there is only a 50:50 chance that the challenge of efficient inference will be met in his lifetime. And that's despite the fact that computers will get faster and algorithms smarter. "These

"The technique can tell within the first 3 hours which babies are likely to be healthy"

problems are much harder than getting to the moon or Mars," he says.

This, however, is not dampening the spirits of the AI community. Daphne Koller of Stanford University, for instance, is attacking very specific problems using probabilistic programming and has much to show for it. Along with neonatologist Anna Penn, also at Stanford, and colleagues, Koller has developed a system called PhysiScore for predicting whether a premature baby will have any health problems – a notoriously difficult task. Doctors are unable to predict this with any certainty, "which is the only thing that matters to the family", says Penn.

PhysiScore takes into account factors such as gestational age and weight at birth, along with real-time data collected in the hours after birth, including heart rate, respiratory rate and oxygen saturation (*Science Translation Medicine*, DOI: 10.1126/scitranslmed.3001304). "We are able to tell within the first 3 hours which babies are likely to be healthy and which are much more likely to suffer severe complications, even if the complications manifest after 2 weeks," says Koller.

"Neonatologists are excited about PhysiScore," says Penn. As a doctor, Penn is especially pleased about the ability of AI systems to deal with hundreds, if not thousands, of variables while making a decision. This could make them even better than their human counterparts. "These tools make sense of signals in the data that we doctors and nurses can't even see," says Penn.

This is why Domingos places such faith in automated medical diagnosis. One of the best known is the Quick Medical Reference, Decision Theoretic (QMR-DT), a Bayesian network which models 600 significant diseases and 4000 related symptoms. Its goal is to infer a probability distribution for diseases given some symptoms. Researchers have fine-tuned the inference algorithms of QMR-DT for specific diseases, and taught it using patients' records. "People have done comparisons of these systems with human doctors and the [systems] tend to win," says Domingos. "Humans are very inconsistent in their judgements, including diagnosis. The only reason these systems aren't more widely used is that doctors don't want to let go of the interesting parts of their jobs."

There are other successes for such techniques in AI, one of the most notable being speech recognition, which has gone from being laughably error-prone to impressively precise (*New Scientist*, 27 April 2006, p26). Doctors can now dictate patient records and speech recognition software turns them into electronic documents, limiting the use of manual transcription. Language translation is also beginning to replicate the success of speech recognition.

Machines that learn

But there are still areas that pose significant challenges. Understanding what a robot's camera is seeing is one. Solving this problem would go a long way towards creating robots that can navigate by themselves.

Besides developing inference algorithms that are flexible and fast, researchers must also improve the ability of AI systems to learn, whether from existing data or from the real world using sensors. Today, most machine learning is done by customised algorithms and carefully constructed data sets, tailored to teach a system to do something specific. "We'd like to have systems that are much more versatile, so that you can put them in the real world, and they learn from a whole range of inputs," says Koller.

The ultimate goal for AI, as always, is to build machines that replicate human intelligence, but in ways that we fully understand. "That could be as far off, and maybe even as dangerous, as finding extra-terrestrial life," says Tenenbaum. "Human-like AI, which is a broader term, has room for modesty. We'd be happy if we could build a vision system which can take a single glance at a scene and tell us what's there – the way a human can." ■

Anil Ananthaswamy is a consultant for *New Scientist*