

# AI robot: evolution of the machines that learn for themselves

In Alan Turing's centenary, no computer can yet pass the 'Turing test' and be taken for human. But, says **Marcus du Sautoy**, the hunt for artificial intelligence is moving in a different, exciting direction that involves creativity, language – and even jazz

I propose to consider the question “Can machines think?” Not my question but the opening of Alan Turing's seminal 1950 paper which is generally regarded as the catalyst for the modern quest to create artificial intelligence. His question was inspired by a book he had been given at the age of 10: *Natural Wonders Every Child Should Know* by Edwin Tenney Brewster. The book was packed with nuggets that fired the young Turing's imagination including the following provocative statement: “Of course the body is a machine. It is vastly complex, many times more complicated than any machine ever made with hands; but still after all a machine. It has been likened to a steam machine. But that was before we knew as much about the way it works as we know now. It really is a gas engine; like the engine of an automobile, a motor boat or a flying machine.”

If the body were a machine, Turing wondered: is it possible to artificially create such a contraption that could think like he did? This year is Turing's centenary so would he be impressed or disappointed at the state of artificial intelligence? Do the extraordinary machines we've built since Turing's paper get close to human intelligence? Can we bypass millions of years of evolution to create something to rival the power of the 1.5kg of grey matter contained between our ears? How do we actually quantify human intelligence to be able to say that we have succeeded in Turing's dream? Or is the search to recreate “us” a red herring? Should we instead be looking to create a new sort of machine intelligence different from our own?

Last year saw one of the major landmarks on the way to creating artificial intelligence. Scientists at

IBM programmed a computer called Watson to compete against the best of America's most successful game shows: *Jeopardy!* It might at first seem the answer quickly enough to beat your rival to the buzzer. This was in fact the final question in the face-off with the two all-time champions of the game show. With the answer “Who is Bram Stoker?” Watson claimed the *Jeopardy!* crown.

Watson is not IBM's first winner. In 1997 IBM's super computer Deep Blue defeated reigning world chess champion Garry Kasparov.



HOW DO YOU SAY SIT? Marcus du Sautoy with a shape-making language robot. BBC

But competing at *Jeopardy!* is a very different test for a computer. Playing chess requires a deep logical analysis of the possible moves that can be made next in the game. Winning at *Jeopardy!* is about understanding a question written in natural language and accessing quickly a huge database to select the most likely answer in as fast a time as possible. The two sorts of intelligence almost seem perpendicular to each other. The intelligence involved in playing chess feels like a vertical sort of intelligence, penetrating deeply into the logical consequences of the game, while *Jeopardy!* requires a horizontal thought process, thinking shallowly but expansively over a large data base.

The program at the heart of Watson's operating system is particularly sophisticated because it learns from its mistakes. The algorithms that select the most likely answers are tweaked by Watson every time it gets an answer wrong so that next time it gets a similar question it has a better chance of getting it right. This idea of machine learning is a powerful new ingredient in artificial intelligence and is creating machines that are quickly doing things that the programmers hadn't planned for.

Despite Watson's win, it did make some very telling mistakes. In the category “US cities” contestants were asked: “Its largest airport is named for a world war two hero; its second largest for a world war two battle.” The humans responded correctly with “Where is Chicago?” Watson went for Toronto, a city that isn't even in the United States.

It's this strange answer that gives away that it is a probably a machine rather than a person answering the question. Getting a machine to pass itself off as human was one of the key hurdles that Turing believed a machine would need to pass in order to successfully claim the realisation of artificial intelligence. With the creation of the Loebner prize in 1991, monetary prizes were offered for anyone who could create a chatbot that judges could not distinguish from the chat of a human being. Called the Turing test, many working in AI regard the challenge as something of a red herring. The Loebner prize, in their opinion, has distorted the quest and has proved a distraction from a more interesting goal: creating machine intelligence that is different from our own.

The AI community is beginning to question whether we should be so obsessed with recreating human intelligence. That intelligence is a product of millions of years of evolution and it is possible that it is something that will be very difficult to reverse engineer without going through a similar process. The emphasis is now shifting towards creating intelligence that is unique to the machine, intelligence that ultimately can be harnessed to amplify our very own unique intelligence.

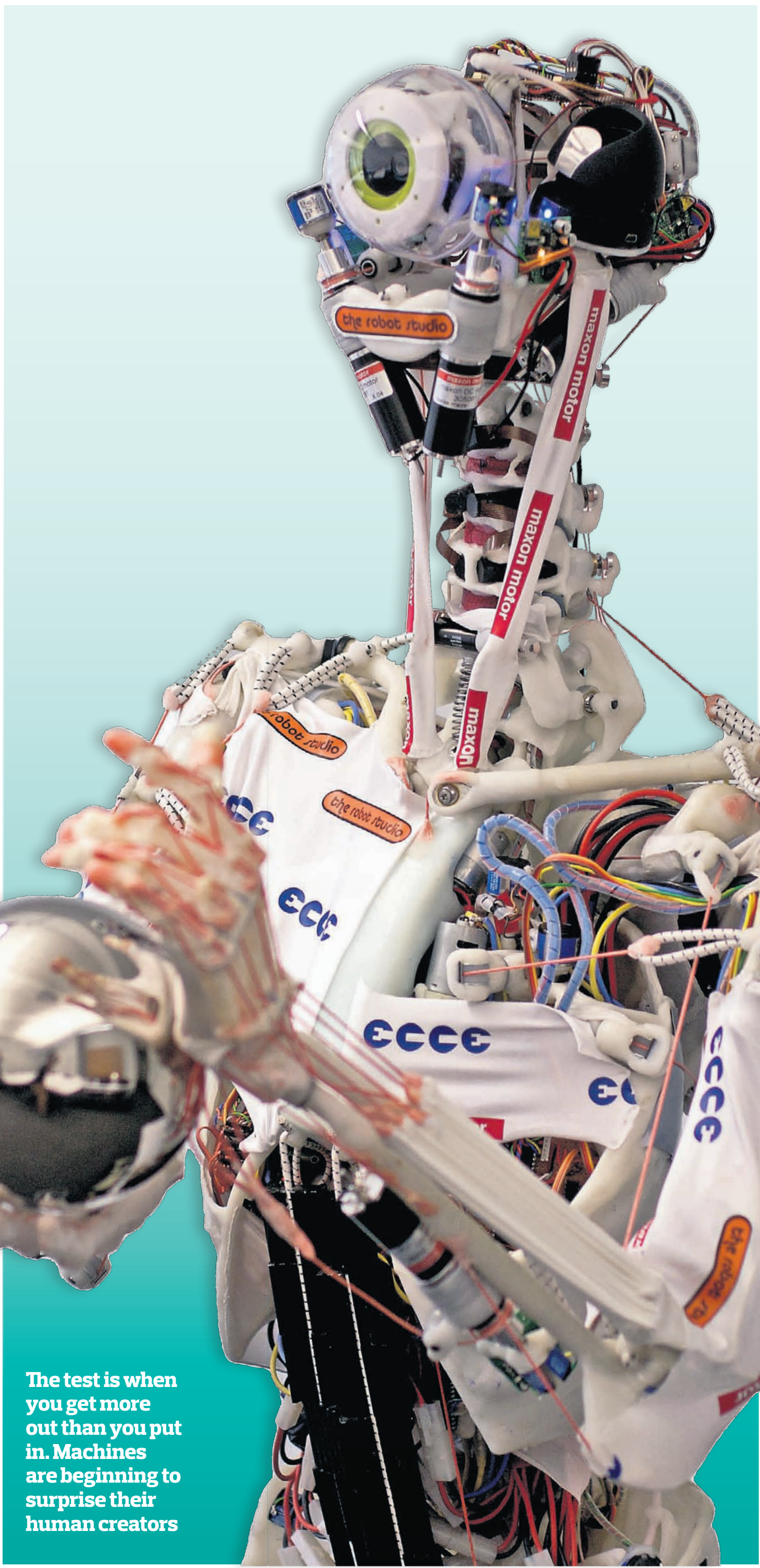
Already the descendants of Deep Blue are performing tasks that no human brain could get anywhere near. Blue Gene can perform 360 trillion operations a second, which compares with the 3 billion instructions per second that an average desktop computer can perform. This extraordinary firepower is being used to simulate the behaviour of molecules at an atomic level to explore how materials age, how turbulence develops in liquids, even the way proteins fold in the body. Protein folding is thought to be crucial to a number of degenerative diseases so these computer simulations could have amazing medical benefits.

But isn't this number-crunching rather than the emergence of a new intelligence? The machine is just performing tasks that have been programmed by the human brain. It may be able to completely outperform my brain in any computational activity but when I'm doing mathematics my brain is doing so much more than just computation. It is working subconsciously, making intuitive leaps. I'm using my imagination to create new pathways which often involve an aesthetic sensibility to arrive at a new mathematical discovery. It is this kind of activity that many of us feel is unique to the human mind and not reproducible by machines.

For me, a test of whether intelligence is beginning to emerge is when you seem to be getting more out than you put in. Machines are human creations yet when what they produce is beginning to surprise the creators then I think you're getting something interesting emerging.

Exciting new research is currently exploring how creative machines can be in music and art. Stravinsky once wrote that he could only be creative by working within strict constraints: “My freedom consists in my moving about within the narrow frame that I have assigned myself for each one of my undertakings.” By understanding the constraints that produce exciting music, computer engineers at Sony's Computer Science Laboratory in Paris are beginning to produce machines that create new and unique forms of musical composition. One of the big successes has been to produce a machine that can do jazz improvisation live with human players. The result has surprised those who have trained for years to achieve such a facility.

Other projects have explored how creative machines can be at producing



**BEING HUMAN** Ecce Robot, a research project that attempts to mimic human movement. The field of AI, however, is moving away from mimicking human thought processes. [www.eccerobot.org](http://www.eccerobot.org)

visual art. The Painting Fool is a computer program written by Simon Colton of Imperial College. Not everyone likes the art produced by the Painting Fool but it would be anaemic art if they did. What's extraordinary is that the programmes in these machines are learning, and changing and evolving so that very soon the programmer no longer has a clear idea of how the results are being achieved and what it is likely to do next. It is this element of getting more out than you put in that represents something approaching emerging intelligence.

For me one of the most striking experiments in AI is the brainchild of the director of the Sony lab in Paris, Luc Steels. He has created machines that can evolve their own language. A population of 20 robots are first placed one by one in front of a mirror and they begin to explore the shapes they can make using their bodies in the mirror. Each time they make a shape they create a new word to denote the shape. For example the robot might choose to name the action of putting the left arm in a horizontal position. Each robot creates its own unique language for its own actions. The really exciting part is when these robots begin to interact with each other. One robot chooses a word from its lexicon and asks another robot to perform the action corresponding to that word. Of course the likelihood is that the second robot hasn't a clue. So it chooses one of its positions as a guess. If they've guessed correctly the first robot confirms this and if not shows the second robot the intended position.

The second robot might have given the action its own name, so it won't yet abandon its choice, but it will update its dictionary to include the first robot's word. As the interactions progress the robots weight their words according to how successful their communication has been, downgrading those words where the interaction failed. The extraordinary thing is that after a week of the robot group interacting with each other a common language tends to emerge. By continually updating and learning, the robots have evolved their own language. It is a language that turns out to be sophisticated enough to include words that represent the concept of “left” and “right”. These words evolve on top of the direct correspondence between word and body position. The fact that there is any convergence at all is exciting but the really striking fact for me is that these robots have a new language that they understand yet the researchers at the end of the week do not comprehend until they too have interacted and decoded the meaning of these new words.

Turing might be disappointed that in his centenary year there are no machines that can pass themselves off as humans but I think that he would be more excited by the new direction artificial intelligence has taken. The AI community is no longer obsessed with reproducing human intelligence, the product of millions of years of evolution, but rather in evolving something new and potentially much more exciting.

Marcus du Sautoy is *Simonyi professor for the public understanding of science and a professor of mathematics at the University of Oxford.*

■ **Horizon: The Hunt for AI**, presented by Marcus du Sautoy, is on BBC2 on Tuesday at 9pm



FREEHAND The Dancing Salesman Problem, created by software the Painting Fool.

## ABSTRACT THOUGHT The robot painter

In 2006 Dr Simon Colton, a researcher in computational creativity at Imperial College, London, started to explore whether a computer program with the capacity to create art could be taken as seriously as a human artist.

Where did the idea for the program you call the Painting Fool come from? As a hobby I wrote software that would turn a photograph into a more artistic piece, but six years ago I brought it into my field of research. I realised that the Painting Fool was a very good mechanism for testing out all sorts of theories, such as what it means for software to be creative. The aim of the project is for the software itself to be taken seriously as a creative artist in its own right, one day.

How does it work? The Painting Fool produces artwork in a number of ways. The first is the simplest one: the software paints according to user-defined input, such as a photograph. But last year I had an exhibition in Paris called No Photos Were Harmed that challenged the public perception of computer art. I presented a couple of pieces, one of which was *The Dancing Salesman Problem* (above), where the figures were generated by a context-free design grammar, which is similar to the grammatical structure of natural language but for images. I've also paired the Painting Fool with emotion-detection software by Maja Pantic, a colleague of mine, so it paints pictures in different styles according to the subject's mood, like the *Really Sad* picture of me (below), where it chose muted colours and graphite pencil. Each of these projects tries to challenge a notion about computer programs – that they can't be imaginative, that they can't appreciate how the output might affect people.

How will you know when it is taken seriously as an artist? People want to know artwork has been constructed with an intelligent thought process, so perhaps once the software produces pieces that are culturally valuable, that get people talking, and are not necessarily anything that I'm keen on aesthetically or conceptually, that would be a good indication of its independence from me.

Interview by Gemma Kappala-Ramsamy

For more information visit [thepaintingfool.com](http://thepaintingfool.com)



## TheObserver reader offer

### Lake Maggiore & the Matterhorn



8 days half-board from only £829pp

Selected departures May to October 2012

Book at [observer.co.uk/holidayoffers](http://observer.co.uk/holidayoffers)  
Call 0330 333 6912

Email: [guardiantraveltravel.co.uk](mailto:guardiantraveltravel.co.uk) Calls to 03 numbers will cost no more than calls to national geographic numbers starting 01 or 02. Prices based on per person sharing a twin room, single rooms and insurance available at a supplement. Holiday organised by Riviera Travel, New Manor, 328 Westmore Road, Burton upon Trent, Staffs DE14 1SP and is offered subject to availability. ABTA V4744 ATOL 3430 protected.

## TheObserver reader offer

### Get your TV ready for the digital switchover

Toshiba DVD/VHS Freeview recorder also transfers VHS & Freeview TV to DVD



In just one machine, and at the touch of a button, you can play VHS video, DVDs, camcorder recordings, or Freeview digital TV and transfer video to DVD (and reverse). Freeview TV to DVD or VHS, camcorder to DVD, or even DVD to DVD using an external player. Comes with a 12-months guarantee and 30 day trial period.

- Features include:
- Built-in Freeview TV receiver with 8-day electronic programme guide – preparing you for when UK TV goes fully digital in April 2012, plus traditional analogue TV tuner
  - Plays DVDs using latest HDMI upscaling (up to 1080p) to your TV for higher definition playback
  - DV-IN for instant high-quality transfer of camcorder recordings to DVD
  - Effectively converts a standard analogue TV into a digital TV
  - Also available with built-in 320GB hard drive (will record over 500 hours of TV): Just £329.95 (save £20 on RRP)

Shop at [guardian.co.uk/offers/t2](http://guardian.co.uk/offers/t2)

or call 0330 333 6805 Quote 0B225

Calls cost the same as dialling a standard 01 or 02 number from your landline or mobile phone. To order by post, please send a cheque to Guardian Reader Offers, (Code 0B225), T2 Direct Ltd, PO BOX 295, Ramsey, SG13 0WH. Delivery within 10 working days. Supply your email address to receive full delivery tracking notifications. Offer subject to availability and free delivery to UK mainland readers only. For special delivery services such as Express, A.M., Saturday and next UK mainland please call our orderline. If you are not completely satisfied, contact us within 30 days of receipt to arrange return of item for a full refund or replacement. All products are purchased from, and your resulting contract will be with T2 Direct Ltd, a company wholly independent of Guardian News & Media Ltd.

Aradio station with Elizabeth I and a keyboard instrument Join them up?

This section, page 44