My New Scientist

Only 4 days remaining to access this article

To ensure full online access subscribe now

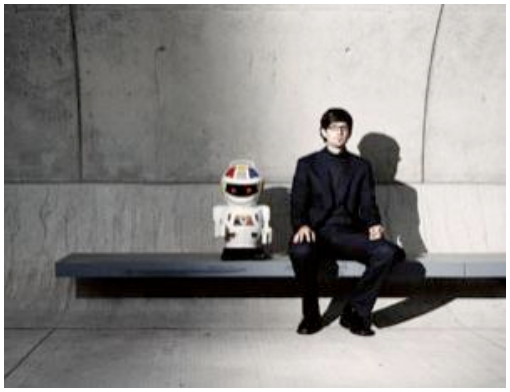
Home | Tech | Life | News | Back to article

Bot shows signs of consciousness

01 April 2011 by **Celeste Biever**
Magazine issue 2806. **Subscribe and save**
For similar stories, visit the **Robots** Topic Guide

Editorial: "When should we give rights to robots?"

A SOFTWARE bot inspired by a popular theory of human consciousness takes the same time as humans to complete simple awareness tasks. Its creators say this feat means we are closer to understanding where consciousness comes from. It also raises the question of whether machines could ever have subjective experiences.



Conscious minds think alike (Image: Erik Dreyer/Stone /Getty)

1 more image

The bot, called LIDA for Learning Intelligent Distribution Agent, is based on "global workspace theory". According to GWT, unconscious processing - the gathering and processing of sights and sounds, for example, is carried out by different, autonomous brain regions working in parallel. We only become conscious of information when it is deemed important enough to be "broadcast" to the global workspace, an assembly of connected neurons that span the brain. We experience this broadcast as consciousness, and it allows information to be shared across different brain regions and acted upon.

Recently, several experiments using electrodes have pinpointed brain activity that might correspond to the conscious broadcast, although how exactly the theory translates into cognition and conscious experience still isn't clear.

To investigate, Stan Franklin, of the University of Memphis in Tennessee, built LIDA - software that incorporates key features of GWT, fleshed out with ideas about how these processes are carried out to produce what he believes to be a reconstruction of cognition.

Franklin based LIDA's processing on a hypothesis that consciousness is composed of a series of millisecond-long cycles, each one split into unconscious and conscious stages. In the first of these stages - unconscious perception - LIDA scans the environment and copies what she detects to her sensory memory. Then specialised "feature detectors" scan sensory memory, pick out certain colours, sounds and movements, and pass these to a software module that recognises them as objects or events. For example, it might discover red pixels and "know" that a red light has been switched on. In the next phase, understanding, which is mainly unconscious, these pieces of data can be strung together and compared with the contents of LIDA's long-term memory. Another set of processes use these comparisons to determine which objects or events are relevant or urgent. For example, if LIDA has been told to look out for a red light, this would be deemed highly salient. If this salience is above a certain threshold, says Franklin, "it suddenly steps over the edge of a cliff; it ignites". That event along with some of its associated content will rise up into consciousness, winning a place in LIDA's global workspace - a part of her "brain" that all other areas can access and learn from. This salient information drives which action is chosen. Then the cycle starts again.

ADVERTISEMENT

Franklin reckons that similar cycles are the "building blocks for human cognition" and conscious experience. Although only one cycle can undergo conscious broadcast at a time, rather like the individual frames of a movie, successive broadcasts could be strung together quickly enough to give the sense of a seamless experience (see diagram).

However, just because these cognitive cycles are consistent with some features of human consciousness doesn't mean this is actually how the human mind works. So, with the help of Baars at the Neuroscience Institute in San Diego, California, who first proposed GWT, and philosophy student Tamas Madl at the University of Vienna, Austria, Franklin put LIDA into direct competition with humans.

To increase her chance of success, they grounded the timings of LIDA's underlying processes on known neurological data. For example, they set LIDA's feature detectors to check sensory memory every 30 milliseconds. According to previous studies, this is the time it takes for a volunteer to recognise which category an image belongs to when it is flashed in front of them.

Next the researchers set LIDA loose on two tasks. The first was a version of a reaction-time test in which you must press a button whenever a light turns green. The researchers planted such a light in LIDA's simulated environment, and provided her with a virtual button. It took her on average 280 milliseconds to "hit" the button after the light turned green. The average reaction time in people is 200 milliseconds, which the researchers say is "comparable".

A second task involved a flashing horizontal line that appears first at the bottom of a computer screen and then moves upwards through 12 different positions. When the rate that it shifts up the screen is slow, people report the line as moving. But speed it up and people seem to see 12 flickering lines. When the researchers created a similar test for LIDA, they found that at higher speeds, she too failed to "perceive" that the line was moving. This occurred at about the same speed as in humans. Both results have been accepted for publication in *PLoS One*.

"You tune the parameters and lo and behold you get that data," says Franklin. "This lends support to our hypothesis that there is a single basic building block for all human cognition." Antonio Chella, a roboticist at the University of Palermo in Italy and editor of the *International Journal of Machine Consciousness* agrees: "This may support LIDA, and GWT as a model that captures some aspects of consciousness."

Murray Shanahan, a cognitive roboticist at Imperial College London who also works on computational models of consciousness, says that LIDA is a "high level" model of the mind that doesn't attempt to represent specific neurons or brain structures. This is in contrast to his own lower-level models, but Shanahan points out that both types are needed: "We don't know what the right theory or right level of abstraction is," he says. "We have to let a thousand flowers bloom."

So is LIDA conscious? "I call LIDA functionally conscious," says Franklin, because she uses a broadcast to drive actions and learning. But he draws the line at ascribing "phenomenal consciousness", or subjectivity to her. That said, he reckons there is no reason in principle why she should not be fully conscious one day. "The architecture is right for supporting phenomenal consciousness if we just knew how to bring it about."

Can a computer ever be aware?

At what point does a model of consciousness itself become conscious - if ever?

Antonio Chella of the University of Palermo, Italy, says that before consciousness can be ascribed to software agent LIDA, (see main story) she needs a body. "Consciousness of ourselves, and the world, is based on a continuous interaction between our brain, our body and the world," he says. "I look forward to a LIDA robot."

However, cognitive roboticist Murray Shanahan at Imperial College London says that the robot need not be physical. "It only makes sense to start talking about consciousness in the context of something that interacts in a purposeful way with a spatio-temporal environment," he says. "But I am perfectly happy to countenance a virtual environment."

LIDA's inventor, Stan Franklin of the University of Memphis, reckons LIDA is already "functionally conscious", but makes a distinction between that and subjectivity or "phenomenal consciousness". He is planning to build a version of LIDA that interacts with humans within a complex environment. "When this happens, I may be tempted to attribute phenomenal consciousness to the agent," he says. The trouble is, even if LIDA could have subjective

experiences, how would we prove it?

We can't even prove that each of us isn't the only self-aware being in a world of zombies. "Philosophers have been dealing with that for more than 2000 years," says Franklin. Perhaps we will simply attribute subjectivity to computers once they become sufficiently intelligent and communicative, he says.

From issue 2806 of New Scientist magazine, page 10-11.

As a subscriber, you have unlimited access to our online archive.

Why not browse past issues of New Scientist magazine?

0 tweets

tweet Like 6

PRINT SEND SHARE

If you would like **to reuse any content** from New Scientist, either in print or online, please **contact the syndication** department first for permission. New Scientist does not own rights to photos, but there are a variety of licensing options available for use of articles and graphics we own the copyright to.

Back to article

PRINT SEND SHARE

ADVERTISEMENT

NewScientist

Bot shows signs of consciousness

A software bot inspired by a theory of human consciousness takes the same time as humans to complete simple awareness tasks

Previous

Add up to awareness

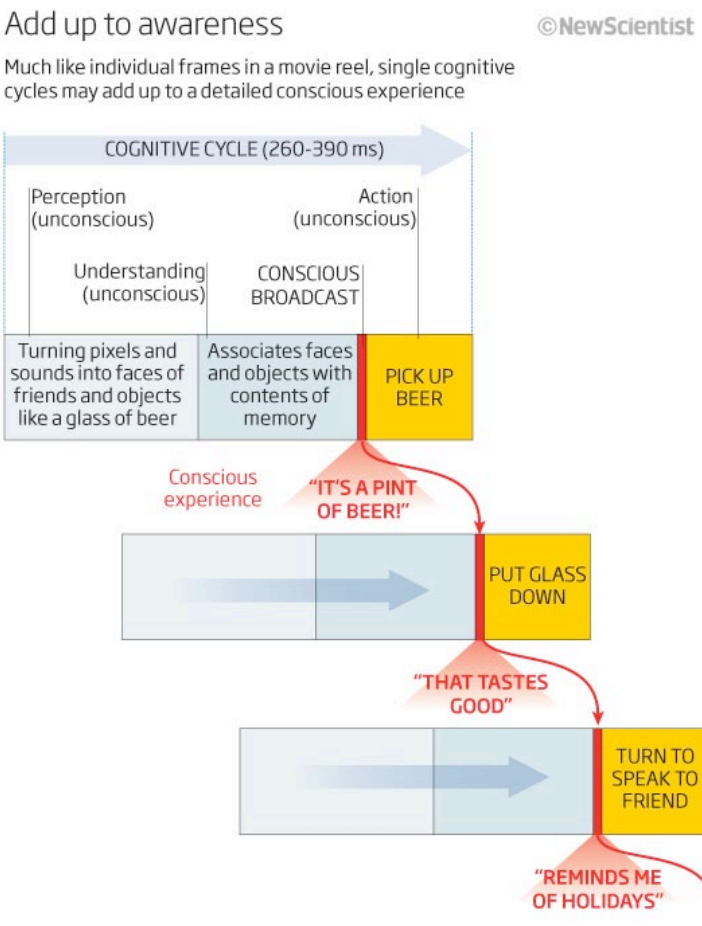


Image 2 of 2