Robots emulating children

Scientists are developing robots using biology as their inspiration. Will they succeed in building cognitive agents?

In many respects, computers are superior to human beings—they can hold more information and easily retrieve every bit of it, they can calculate the square root of any number within a fraction of a second, and they even beat humans at chess. However, there are many tasks in which computers do not stand a chance against a human counterpart, even a toddler. Take, for example, ASIMO, one of the most sophisticated humanoid robots, developed by Honda (Tokyo, Japan). ASIMO can walk around a party, shake hands with the guests and serve food. However, it still walks relatively clumsily and slowly, and can carry out only a limited number of pre-programmed actions—trying to engage ASIMO in a conversation about politics would be futile. AIBO, the robot dog produced by Sony (Tokyo, Japan), seemed promising in the beginning; however, it barked up the wrong tree too often and did not behave like a real dog, so Sony eventually cancelled its production. Such is the disappointing state of affairs in the creation of robots with ‘intelligent’ behaviour. Is there a chance that future robots will fare any better?

In fact, there is. Robot scientists are turning to new strategies that are inspired by neuroscience. This is a major step away from earlier approaches in which any potential task had to be programmed explicitly—for every new task that a computer or robot had to perform, the programme behind it had to be adapted. This can actually work successfully, as long as the computers perform according to strict rules in a static environment, such as when playing chess. However, succeeding in a more complex environment with unexpected challenges requires far more flexibility than any hardwired programme can achieve. The behaviour of humans and

animals in everyday life is far too complex to ever be formulated in any programming language—even a seemingly simple task, such as the coordination of movements, has constantly challenged robot scientists. Moreover, no engineer can predict every situation that a robot might encounter, and programme its reactions accordingly. Therefore, for a robot to perform in a natural environment, it must be able to adapt its behaviour autonomously. In short, it needs to learn—just as animals and humans do—and this is exactly where robotics is heading.

These new generation computers and robots imitate ‘cognitive’ behaviour by emulating various aspects of human or animal learning. “The human brain is a good model for robot learning. There is nothing in nature that learns more efficiently,” said Florentin Wörgötter, Professor of Informatics at the Bernstein Centre for Computational Neuroscience and the University of Göttingen (Germany). This raises the question to what extent neural processes can be transferred to silicon. Will such robots develop reasoning and autonomously find strategies to solve problems? Will they outperform our expectations and act more autonomously than we want them to?

Machine learning can be achieved at different levels of complexity, much as different scientific fields—from cellular biology to aetiology—investigate learning processes at different levels of biological complexity. At the most basic cellular level, Hebbian learning describes the mechanism that increases synaptic efficacy when the pre-synaptic and post-synaptic neurons fire in short sequence, one after the other—“what fires together, wires together”. Transferring such rules into mathematical equations to emulate learning processes is actually relatively simple (Arbib, 1995; Torras, 2002).

However, Hebbian learning is not practical for programming more complex behaviour—it would amount to reconstructing

the brain by simulating every synapse required for any expected behaviour. To simulate more complex learning, scientists have therefore taken a step up the ladder of complexity. Reinforcement learning, for example, is a mathematical model frequently used in artificial intelligence and cognitive science (Torras, 2002), which parallels the brain’s dopamine-based reward system.

...the question remains as to whether we want robots to acquire too much autonomy

Dopamine-releasing neurons perform an important function in learning behavioural reactions controlled by reward—they encode rewarding aspects of environmental stimuli like food, sex or pleasure. Neurons with different response types are important for distinct aspects of reward learning. For example, ‘prediction-error neurons’ increase their firing rate in response to a reward-predicting stimulus. The mathematical model of reinforcement learning reflects the dopamine-based reward system down to the level of single cells: specific characteristics of prediction-error and other neurons of the dopamine system have their mathematical counterparts in this model (Montague *et al.*, 2004; Wörgötter & Porr, 2005). Reinforcement-learning algorithms therefore prompt the robot to carry out behaviours that maximise whatever it has been programmed to regard as reward in much the same way as the dopaminergic system does in animals. Just like reinforcement learning, many other mathematical models now induce various forms of learning with strong parallels to higher organisms. Teaching robots to learn has already proved to be successful in various applications. For instance, it allows robots to navigate through unknown terrain without bumping into walls or to improve coordination, making them move more smoothly and efficiently.

...for a robot to perform in a natural environment, it must be able to adapt its behaviour autonomously

There is more to come. Fed with the right algorithms, future robots might be able to acquire much more sophisticated skills than just being able to navigate through a maze. One such ability would be to ‘understand’ the function of an object, which is something that babies and toddlers must also learn. Although today’s robots can ‘see’ objects and handle them, they still lack any understanding of their essence or function. A cup, for example, can be filled with liquid and used to drink from—be it a mug, a tea cup or an espresso cup. However, a robot is still not able to classify these objects as being different in shape, size and colour, yet belonging to the same category based on their function.

Programming a robot with curiosity might help it to acquire such abilities. “In principle, this can be done and has already been attempted in some labs,” said Wörgötter. Curiosity could be implemented by programming the robot to regard novelty as reward. Such a robot would learn in the same way as a baby or a toddler—by showing a keen interest in and exploring the environment to investigate anything that is unknown. A curious robot might eventually be able to understand the function of objects and acquire cognition. Although science is still far from creating such robots, their construction is feasible. “The basic concepts required to construct cognitive robots are probably already there. I don’t think it will require an Einstein of cognitive

science, who will suddenly understand brain function and put it into a single formula,” said Wörgötter. “It is more a matter of bringing together multiple components into an autonomous dynamic system.”

If free will requires no ‘soul’ but can emerge from physical matter, such as the brain, why should it not also emerge from silicon?

Several joint research projects are now underway across the European Union to develop robots with cognitive abilities that can interact with humans in a more sophisticated manner. Researchers in the field are quite optimistic about the feasibility of this work based on algorithms that emulate various aspects of human learning. According to the web site for PACO-PLUS—a project that aims to design a robot combining perception, action and cognition—robots “need to understand both what they perceive and what they do” (www.paco-plus.aau.dk). The cognitive robot companion COGNIRON “is not only considered as a ready-made device but as an artificial creature, which improves its capabilities in a continuous process of acquiring new knowledge and skills” (www.cogniron.org). According to the project description of MirrorBot, “...a biological and neuroscience-oriented approach for multimodal processing will lead to new

life-like perception action systems” (www.his.sunderland.ac.uk/mirrorbot).

“The cognitive abilities of humans are very complex and include a multimodal integration of different cognitive processes. In the MirrorBot project we consider different cognitive processes together,” explained Stefan Wermter, Professor of Computer Science at the University of Sunderland (UK), and Coordinator of MirrorBot. The aim is to construct robots that ‘understand’ and can relate their actions to spoken language, and will be able to accomplish simple tasks that they are instructed to perform, such as selecting certain objects from a table and grasping them. The project is inspired by the concept of mirror neurons (Wermter *et al*, 2005)—nerve cells that fire both when an animal perceives an action, through sound or vision, and when it carries out the same action. This indicates that these neurons have a role in linking perception and action. Mirror neurons are also involved in understanding language. “The robot can then learn through instructions,” said Wermter, which would be another leap forward in the implementation of learning in robots.

However, to what extent the combination of cognitive sciences and robotics will succeed in building more human-like robots is not clear. Moreover, the question remains as to whether we want robots to acquire too much autonomy. “We don’t want a service robot to get so autonomous that it decides to

take a rest rather than vacuuming the floor," said Wörgötter. Although he said this jokingly, the research does raise interesting questions of both biology and philosophy: Can a machine, an artificial construct, make decisions? Will it eventually acquire 'free will'?

If scientists ever do construct human-like robots, we will have great difficulties in understanding 'what'—or 'who'—we are dealing with

There are different interpretations of what human free will is and where it comes from (Greene & Cohen, 2004; Rose, 2005). As our understanding of the brain grows, the old view of an immaterial 'mind' or 'soul' that is responsible for complex neuronal functions, such as decision making, becomes increasingly outdated (Farah, 2005). Instead, the view prevails that free will emerges from physical processes in the brain—an idea that is often referred to as 'compatibilism'. If free will requires no 'soul' but can emerge from physical matter, such as the brain, why should it not also emerge from silicon? Indeed, any simple computer can make 'decisions'—it only needs information and algorithms to make a choice. Simple systems that make such 'decisions' already exist in great numbers. A thermostat 'decides' to turn on the heating system if the room temperature drops below a certain level. "It certainly doesn't think in the way humans do, but the question is whether that is just a different degree of complexity, or whether there is some really qualitatively different process that makes human choice not just a physical process," said Joshua Greene, a psychologist at Princeton University (NJ, USA). According to Greene, it is a matter of complexity: "Our conception of ourselves as being above the laws of nature is an illusion."

"We don't know whether we can make something out of silicon that is intelligent, but I think, and so do other people, that

we probably can in the end," said Ned Block, Professor of Philosophy and Psychology at New York University (NY, USA). "My view is that 'free will' will just come with intelligence." The possibility of constructing such robots strengthens the notion that free will in humans is a matter of physical processes—an idea that many people are still reluctant to accept. "The concept of the soul might become obsolete," said Wörgötter.

However, it will not be straightforward to determine whether a robot actually has intelligence that is sufficiently genuine to produce free will (Block, 1995; Buttazzo, 2001; Franklin, 1995). A seemingly intelligent being can be created using simple algorithms that simulate some aspects of human behaviour. For example, in 1966, Joseph Weizenbaum, then at the Massachusetts Institute of Technology (Boston, MA, USA), developed a 'computer therapist' called ELIZA, which analyses written sentences by detecting keywords, and formulates an appropriate response according to simple rules, despite having no real understanding of what it is saying. Nonetheless, ELIZA fooled some people into thinking that they were communicating with a fellow human (Block, 1995). "Whether a robot is intelligent or not is a matter not just of its behaviour but how the behaviour is produced," said Block. Systems that only fool us will be debunked sooner or later. "I think when people come up with a computational structure for an intelligent robot, we will very well agree that it is intelligent," said Block.

However, what is true for intelligence might not be true for consciousness (Buttazzo, 2001). "There is more of a problem in figuring out whether a robot is conscious. We would need to understand not just the biology of human consciousness, but also consciousness in general," commented Block. "We would need some kind of conceptual and theoretical breakthrough to be able to tell whether a robot is conscious."

If scientists ever do construct human-like robots, we will have great difficulties in understanding 'what'—or 'who'—we are dealing with. First of all, we might find it hard to evaluate whether a robot has a free will that is based on understanding. In this context, the consequences are nebulous. If it has a free will, would we need robot laws, as suggested by the science fiction writer Isaac Asimov? Moreover, we will probably never know whether a robot has self-awareness or consciousness. Maybe consciousness requires something that is unique to the human brain and cannot be reproduced in silicon, but there will be no way to judge. How will we relate to such an object?

REFERENCES

- Arbib MA (1995) *Handbook of Brain Theory and Neural Networks*. Cambridge, MA, USA: MIT Press
- Block N (1995) The mind as the software of the brain. In Osherson D, Gleitman L, Kosslyn S, Smith E, Sternberg S (eds) *An Invitation to Cognitive Science*, Vol. 3. Cambridge, MA, USA: MIT Press
- Buttazzo G (2001) Artificial consciousness: utopia or real possibility? *Computer Mag* **34**: 24–30
- Farah MJ (2005) Neuroethics: the practical and the philosophical. *Trends Cogn Sci* **9**: 34–40
- Franklin S (1995) *Artificial Minds*. Cambridge, MA, USA: MIT Press
- Greene J, Cohen J (2004) For the law, neuroscience changes nothing and everything. *Philos Trans R Soc Lond B Biol Sci* **359**: 1775–1785
- Montague PR, Hyman SE, Cohen JD (2004) Computational roles for dopamine in behavioural control. *Nature* **431**: 760–767
- Rose SP (2005) Human agency in the neurocentric age. *EMBO Rep* **6**: 1001–1005
- Torrás C (2002) Neural computing increases robot adaptivity. *Natural Comput* **1**: 391–425
- Wermter S, Palm G, Elshaw M (eds; 2005) *Biometric Neural Learning for Intelligent Robots: Intelligent Systems, Cognitive Robotics, and Neuroscience*. Heidelberg, Germany: Springer
- Wörgötter F, Porr B (2005) Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. *Neural Comput* **17**: 245–319

Katrin Weigmann

doi:10.1038/sj.embo.7400694