

Hard hearted? In *I, Robot*, Will Smith finds there is more to the robots than meets the eye.

## Robots have feelings too

### Who Needs Emotions? The Brain Meets the Robot

edited by Jean-Marc Fellous & Michael A. Arbib

Oxford University Press: 2005. 399 pp. £36.50, \$59.95

#### Dylan Evans

As with so much else in the field of artificial intelligence, the idea that robots might one day have emotions first appeared in science fiction. The Czech playwright Karel Capek, who coined the word 'robot' in 1921, pictured robots rebelling against their creators. Isaac Asimov went on to imagine robots with more positive emotions. Philosophers soon got in on the act and within a decade had explored many of the conceptual puzzles posed by these stories. With one or two notable exceptions, however, it wasn't until the 1990s that scientists and engineers finally began the attempt to turn the stories into reality.

Broadly speaking, research in emotional robotics can be divided into two distinct approaches. Some researchers prefer to concentrate on the practical task of giving robots the ability to interact with humans in emotional ways, such as detecting emotional states in people, or behaving in ways that are readily interpreted by people as expressions of emotion. Others set themselves the more ambitious task of endowing robots with an artificial analogue of the emotional-motivational system common to humans and many other animals. The two approaches reflect different goals: the first aims simply to produce robots that can interact socially with humans, whereas the second aims to deepen our understanding of what emotions really are.

*Who Needs Emotions?* deals mainly with the second approach. The book is a collection of

essays by some of the leading lights in affective neuroscience and emotional robotics. The editors, Jean-Marc Fellous and Michael Arbib, begin the book with a short dialogue between two imaginary characters, one with a theoretical bent, the other with a resolutely practical approach. This device allows them to introduce some of the key difficulties that beset the field in a remarkably concise and balanced manner. In just five pages, they give a glimpse of the thorny problems surrounding the definition of emotional terms and the difficulty of knowing when such terms apply to robots.

The two imaginary characters also serve as a way of classifying the other contributions to the book. Most of them are theoretical, with only a few describing actual attempts to build emotional robots. In the latter category, Ronald Arkin reviews a variety of simple behaviour-based robots that provide well defined models for a number of simple drives such as hunger, fear and sex, and Cynthia Breazeal and Rodney Brooks describe Kismet, an emotionally expressive robot that they designed to interact with people in a lifelike way. These are the most interesting chapters, because the existence of a real system, a concrete bit of hardware, allows for detailed questions to be asked about the differences and similarities between artificial and natural emotions.

The theoretical chapters present a variety of different models of emotion that aim both to describe how emotions work in biological creatures such as humans, and to serve as a blueprint for implementing emotions in robots. But these models lack any hardware implementation, so they are not easy to evaluate. In fact it is often difficult to see how they might be implemented, sometimes because they could be implemented in many different ways, and sometimes because they are simply

confusing. In the H-CogAff model proposed by Aaron Sloman, Ron Chrisley and Matthias Scheutz, for example, there are so many arrows in the diagram that it seems that every subsystem is connected with every other one, which calls into question the value of their distinction between reactive mechanisms, deliberative reasoning and meta-management. Andrew Ortony, Donald Norman and William Revelle also outline a tripartite model in which emotions are considered at three different levels of information processing — the reactive, the routine and the reflective — but these are not easily mapped on to the categories set out by Sloman, Chrisley and Scheutz.

This lack of theoretical agreement is characteristic of the whole field. Every researcher in this area has a pet theory, and yet there seems to be no agreement even as to the criteria that would allow us to choose one theory over another, let alone as to which theory is best. If some degree of consensus is a sign of scientific maturity, emotional robotics is still very much in its infancy.

But this is precisely what makes it such a difficult and interesting area of study. If you enjoy the kind of research that the philosopher of science Thomas Kuhn called "normal science", in which the ground rules are long established and it is now simply a question of solving well defined problems, this volume is not for you. But if you prefer the more disorienting and less predictable task of discovering the ground rules themselves, these essays on emotional robotics provide a sometimes frustrating, but always fascinating, glimpse of a science in the making. ■

Dylan Evans is a senior lecturer in intelligent autonomous systems at the University of the West of England, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK.