On our best behaviour

Hector J. Levesque
Dept. of Computer Science
University of Toronto
Toronto, Ontario
Canada M5S 3A6
hector@cs.toronto.edu

Abstract

The science of AI is concerned with the study of intelligent forms of behaviour in computational terms. But what does it tell us when a good semblance of a behaviour can be achieved using cheap tricks that seem to have little to do with what we intuitively imagine intelligence to be? Are these intuitions wrong, and is intelligence really just a bag of tricks? Or are the philosophers right, and is a behavioural understanding of intelligence simply too weak? I think both of these are wrong. I suggest in the context of question-answering that what matters when it comes to the science of AI is not a good semblance of intelligent behaviour at all, but the behaviour itself, what it depends on, and how it can be achieved. I go on to discuss two major hurdles that I believe will need to be cleared.

1 Intelligent behaviour

This paper[1] is about the *science* of AI. Unfortunately, it is the *technology* of AI that gets all the attention. The general public could be forgiven for thinking that AI is just about all those whiz-bang applications, smart *this* and autonomous *that*. Those of us in the field know that for many applications, the term "intelligent" is no more than a buzzword (like the term "delicious" in "red delicious apples"). And along with the many possibly beneficial AI applications under consideration, we often have serious misgivings about the potential misuse of AI technology (in areas like weaponry).

But AI is more than just technology. Many of us are motivated not by any of the AI applications currently being considered, but by the scientific enterprise, the attempt to understand the world around us. Different sciences have different subject matters, and AI is the study of *intelligent behaviour* in computational terms. What could be more fascinating? The human brain is a remarkable thing, perhaps the single most complex object we know of in the universe. But even more remarkable is what a human brain is capable of *doing*. Our intelligent behaviour at its best goes well beyond what we have any right to expect to emerge out of purely physical matter. Indeed, the overarching question for the science of AI is:

> How is it possible for something physical (like people, for instance) to actually do X?

where X is one of the many instances of intelligent behaviour. This needs to be contrasted with a related question:

> Can we engineer a computer system to do something that is vaguely X-ish?

about which we will have much more to say later.

Note that the science of AI studies intelligent behaviour, not *who* or *what* is producing the behaviour. It studies natural language understanding, for instance, not natural language understanders. This is what makes AI quite different from the study of *people* (in neuroscience, psychology, cognitive science, evolutionary biology, and so on).

What sort of behaviour do we care about? Different researchers will quite naturally focus on different aspects. The behaviour may or may not depend on perceptual or motor skills. It may or may not include learning. It may or may not be grounded in emotional responses, or in social interactions. For some researchers, the main concern is intelligent behaviour seen in a variety of animals, like the ability to find a desired object in a room. For others, the focus is on behaviour seen in humans only, like the ability to play chess. (These two groups sometimes engage in methodological disputes, with the former arguing that we cannot expect to understand human behaviour until we understand its more basic forms, and the latter responding that this is not how science works at all. At this stage of the game, there is really no reason to take a doctrinaire position one way or another.)

1.1 Answering questions

In this paper, I intend to examine one basic form of intelligent behaviour: answering certain *ad-hoc* questions posed in English. Consider a question like the following:

> *Could a crocodile run a steeplechase?*

Even if you know what crocodiles and steeplechases are,[2] you have never really thought about this question before, unless you happened to have read an early paper of mine [6]. Nor

[1] This paper is a written version of the Research Excellence Lecture presented in Beijing at the IJCAI-13 conference. Thanks to Vaishak Belle and Ernie Davis for helpful comments.

[2] For those who do not know, a steeplechase is a horse race, similar to the usual ones, but where the horses must jump over a number of hedges on the racetrack. So it is like hurdles for horses.