Lecture 16: Central Themes

s.butterfill@warwick.ac.uk

1. The Question

What is the mark that distinguishes actions? (Davidson 1971).

'The problem of action is to explicate the contrast between what an agent does and what merely happens to him' (Frankfurt 1978, p. 157).

2. Questions We Couldn't Answer

[Frankfurt] Do spiders have intentions?

[Bach] When I duck to avoid a flying object, am I acting on any intention?

There are theoretically coherent ways of characterising intention on which the answers are yes.

There are theoretically coherent ways of characterising intention on which the answers are no.

How can we decide between these competing views of intention, given that both are theoretically coherent?

3. Are scientific discoveries relevant?

Donald Davidson asks, 'What is the mark that distinguishes … actions?' Are scientific discoveries relevant to answering this question?

3.1. Aristotelian Approaches to Action

What mundane purposes does thinking about actions serve?

- prediction and coordination
- ethical (assigning responsibility, blame; living together)
- normative (how actions should be)
- regulative (he wants himself and others to live it out as much as to describe how things are)

Insofar as thinking about actions enables making predictions, how accurate would we expect it to be?

Whenever you are making predictions about anything at all, you face a **trade-off between accuracy and speed**. Making more accurate predictions requires considering more information and integrating it in a more complex model of minds and actions. By contrast, making faster predictions requires narrowing the information you consider and using a less complex model of minds and actions. Since an observer often has to make predictions fast enough to actually coordinate her actions with another agent's, and since making predictions consumes scarce cognitive resources, the observer usually needs to trade accuracy for speed.

Because making predictions involves a trade-off between speed and accuracy, we should not expect mundane thinking about actions to be especially accurate.

Relying on philosophers to characterise actions would be like relying on Aristotelians to characterise physical objects.

4. Two Kinds of Motivational State

Preferences are one kind of motivational state. These states change under the influence of data-driven learning and fashion, among other things. I prefer chocolate over rhubarb right now, but might later have the reverse preference.

Another kind of motivational state is what animal learning theorists call 'primary motivational states'. These are not modifiable by data-driven learning (nor fashion), or at least not readily modifiable. They include hunger, thirst, lust and disgust.

Can your primary motivational states diverge from your preferences?

4.1. Premises

1. Toxicosis directly influences only primary motivational states.

2. Primary motivational states directly influence only stimulus-driven actions.

To a first approximation, the *stimulus-driven* actions are those actions formed in the presence of stimuli because of the stimuli's presence (not driven by representations of the stimulus).

4.2. Devaluation - standard procedure

- Training: Rat is put in chamber with Lever; pressing Lever dispenses sucrose (novel food).
- Devaluation: Rat is taken into another chamber, poisoned, and then exposed to sucrose.
- Extinction Test: Rat returns to chamber with Lever; pressing Lever does nothing.

4.3. Dissociating preferences from primary motivational states

In rats (and humans), we can dissociate at least two kinds of states involved in causing actions. Namely, preferences and primary motivational states.

These are linked to distinct processes, and distinct patterns of explanation (both of which involve justifying reasons of one or another kind).

Intentions (as well as beliefs and desires) play a role in one pattern of explanation, but not in the other.

'The pattern of results accords [...] with a role for an incentive learning process in the reinforcer devaluation effect; not only must consumption of the reinforcer be paired with toxicosis, the animals must also have an opportunity to contact the reinforcer after aversion conditioning if there is to be a change in instrumental performance' (Balleine & Dickinson 1991, p. 293)

'primary motivational states, such as hunger, do not determine the value of an instrumental goal directly; rather, animals have to learn about the value of a commodity in a particular motivational state through direct experience with it in that state' (Dickinson & Balleine 1994, p. 7)

'primary motivational states have no direct impact on the current value that an agent assigns to a past outcome of an instrumental action; rather, it appears that agents have to learn about the value of an outcome through direct experience with it, a process that we refer to as *incentive learning*' (Dickinson & Balleine 1994, p. 8)

4.4. Steve's objection from primary motivational states

This argument depends on the finding that your primary motivational states can diverge from your preferences, and the inferences listed in Section 4.3.

1. In entering the magazine, our rat is acting.
2. The rat's action is driven by a primary motivational state and a stimulus (the sucrose solution).

Therefore (from 2):

3. The rat's action does not involve intention.

Therefore (from 1 and 3):

4. Not all actions involve intentions.

5. Conclusion

Different marks distinguish different kinds of action. To find the marks, identify the patterns of explanation.

References

Balleine, B. & Dickinson, A. (1991). Instrumental performance following reinforcer devaluation depends upon incentive learning. *The Quarterly Journal of Experimental Psychology Section B*, *43*(3), 279–296.

Davidson, D. (1971). Agency. In R. Binkley, R. Bronaugh, & A. Marras (Eds.), *Agent, Action, and Reason*, (pp. 3–25). Toronto: University of Toronto Press. Reprinted in Davidson, D. (1980) *Essays on Actions and Events*. Oxford: Oxford University Press.

Dickinson, A. & Balleine, B. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior*, *22*(1), 1–18.

Frankfurt, H. (1978). The problem of action. *American Philosophical Quarterly*, *15*(2), 157–162.