Interacting Mindreaders

Stephen A. Butterfill
<s.butterfill@warwick.ac.uk>

November 25, 2011

Abstract

***

1. Goal ascription

Purposive action is action directed to the realisation of one or more outcomes. Identifying which outcomes others' purposive actions are directed to enables one to predict and manipulate their actions, and to emulate their achievements. It is also instrumental for mindreading: knowing which outcomes an action is directed to may constrain hypotheses about what an agent intends as well as potentially providing information concerning what the agent knows, believes or desires. For example, if we know that the goal of an agent's action is to retrieve the key to her house, and if we also know where her key is, we may be able to infer that she does not know where her key is, or that she falsely believes that her key is in her purse.

***Examples.

These simple facts about goal ascription raise many questions. Some concern mechanism, how in fact one subject is able to discover facts about which outcomes another agent's actions are directed to. Another set of questions focuses on the evolution of goal ascription and the costs and benefits of being able to ascribe goals and of being a potential target of goal ascription. Our concern here is not directly with any of these questions. Instead we shall focus on a more narrowly epistemic question. What evidence could support hypotheses about the outcomes to which actions are directed? And how would the evidence support the hypotheses?[1]

---

[1] These questions are versions of those Davidson constructs a theory of interpretation to answer. While what follows draws freely on Davidson's insights, our aims here are more modest than his. For we are concerned only with a fraction of the problem of ascribing mental states and meanings; and, unlike Davidson, we are not concerned with larger claims about the nature of mind. See Davidson (1973, 1990); Lepore & Ludwig (2005).

Of special interest is evidence available independently of any knowledge of mind or language. We want to know how it is possible to identify goals even without knowing what an agent believes or desires and even without understanding their communicative actions. Accordingly we will adopt the perspective of a goal ascriber who knows nothing about the mental states of her target agent that would distinguish this agent from any other. We will also stipulate that there is initially no common ground, shared culture or conventions. And we will stipulate that the goal ascriber is initially unable to understand any communicative actions.

There are two sorts of motivation for these restriction on the evidential basis. One is simply that developmental and comparative research indicates that goal ascription does appear to take place in such circumstances. This makes it important to understand the grounds on which such ascriptions may be based (this may not tell us very much about how goal ascriptions are actually made, but it will be necessary to understand reliable success of mechanisms for goal ascription). Another source of motivation is the conjecture that goal ascription is a prerequisite for the more sophisticated mindreading activities which reveal mental states and meanings. The coherence of this conjecture depends on the possibility of knowing something about which outcomes an agent's actions are directed to independently of knowing what she believes or desires and independently of understanding her communicative actions.[2]

So what evidence could support goal ascription by someone who knows nothing discriminating about her targets' mental states or communicative actions? The standard view, simplified, is this.[3] Faced with an action, the would-be goal ascriber first asks which outcomes this action could be a means to realising. She then considers which of these outcomes are potentially beneficial for, or desirable to, the agent. Any such outcomes are identified as

---

[2] *Compare and contrast Davidson? He did think relational attitudes (holding true) are the foundation for interpretation. But he also thought that interpretation had to happen all at once.)

Bennett (1976, pp.48–50) suggests that a theory of goals ascription has to be developed together with a theory of (proto-)belief ascription: 'An animal's behaviour does not show what it registers unless we know what it seeks; but how can we learn what it seeks before we know what it registers?' [p. 48]

Dennett (1987, p. 17) 'Here is how it works: first you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally your predict that this rational agent will act to further its goals in the light of its beliefs. A little practical reasoning from the chosen set of beliefs and desires will in many—but not in all—instances yield a decision about what the agent ought to do; that is what you predict the agent will do.'

[3] *ref? Dennett?

goals to which the action is directed. So the fact that an action is a means to realising some outcome which is potentially beneficial or desirable is evidence for the conclusion that this outcome is one to which the action is directed.

(E$_1$) Action a is a means to realising outcome G.

(E$_2$) The occurrence of outcome G is potentially beneficial for, or desirable to, the agent of a.

Are collectively evidence for concluding that:

(C) G is a goal to which action a is directed.

References

Bennett, J. (1976). *Linguistic Behaviour.* Cambridge: Cambridge University Press.

Davidson, D. ([1984] 1973). Radical interpretation. In *Inquiries into Truth and Interpretation* (pp. 125–139). Oxford: Oxford University Press.

Davidson, D. (1990). The structure and content of truth. *The Journal of Philosophy, 87*(6), 279–328.

Dennett, D. (1987). *The Intentional Stance.* Cambridge, Mass.: MIT Press.

Lepore, E. & Ludwig, K. (2005). *Donald Davidson: Meaning, Truth, Language, and Reality.* Oxford University Press.