

Mindreading and Joint Action: Philosophical Tools Notes for Lecture 3

Stephen A. Butterfill
<ButterfillS@ceu.hu>

October 23, 2012

1. Introduction

This lecture is going to be a tutorial on some topics in the philosophy of action. Why? We're at a cross-roads between mindreading and joint action.

Looking back to mindreading, accounts of mindreading often focus on belief and desire while completely neglecting action and intention. Outside of linguistic contexts, when you ascribe mental states you often do so on the basis of observing actions. But what are these actions? Should we think of them as mere movements, just like the movements of any other bodies. No. Starting with mere movements doesn't make sense because beliefs and desires are characterised partly in terms of their consequences for action, as we saw. So mindreading needs to involve understanding action, not just understanding mental states. And in line with my general strategy, I suggest that to understand what it is to understand action it is helpful first to ask what action is.

A second reason arising from mindreading is that the focus on belief and desire results in us ignoring mental states like intention. Since at least some forms of mindreading involve ascribing intentions, not just beliefs and desires, it is worth asking what distinguishes intentions from other mental states like beliefs and desires.

Looking forward to our next topic, joint action, we also find reasons for wanting to study action and intention. Joint actions are actions involving multiple agents. To think seriously about joint action it will be helpful to have some background on what action and intention generally.

So that's why we're shifting into tutorial mode today.

2. Question

One feature of actions is that they are directed to goals. I seize little Isabel by the wrists and swing her around, thereby making her laugh and breaking a vase. You might wonder what the goal of my action was. Did I act in order to break the vase or to make Isabel laugh? Or was my action perhaps directed to some other goal, one that was not realised so that my action failed?

A basic challenge for an account of action is to explain the relation between actions and the goal or goals to which they are directed. Among all of the actual and possible outcomes of an action, which are goals of the action?

The standard answer to this question involves intention. An intention (1) represents an outcome, (2) causes an event; and (3) causes an event whose occurrence would normally lead to the outcome's occurrence. What singles out an actual or possible outcome as one to which the component actions are collectively directed? It is the fact that this outcome is represented by the intention.

So the intention is what links an action to the outcomes to which it is directed.

3. Goals are not intentions

Now we have to ask, What is an intention?

First note that goals are not intentions. Goals are actual or possible outcomes. They are states of affairs. Intentions, by contrast, are or involve mental states that represent a goal. It would be a terrible mistake to confuse a goal with a goal-state. That would be like confusing a person with a photograph. ***

4. What is intention?

The recap, the question is what links an action to the outcome or outcomes to which it is directed. And the standard answer is that an intention does this. Intentions are what link actions to the outcomes to which they are directed.

Now we have to ask, What is an intention?

The first idea to consider is, surprisingly, that there are no such things as intentions.

The expression 'the intention with which James went to church' has the outward form of a description, but in fact it ... cannot be taken to refer to an entity, state, disposition, or event. Its function in context is to generate new descriptions of actions in terms of their reasons; thus 'James went to church with the intention

of pleasing his mother' yields a new, and fuller, description of the action described in 'James went to church.' (Davidson 1963, p. 690)

What motivates this view? We already have beliefs and desires in our model of action explanation. Introducing intentions as additional mental states would make the model more complicated. So if we can do without intentions, we should do so in the interests of simplicity.

But how can we do without intentions? Haven't we just seen that we need intentions in order to explain the relation between an action and the goal or goals to which it is directed?

Here's how Davidson's view works. James desired to please his mother. James believed that going to church would please his mother. And this belief and desire caused his going to church.

So the belief-desire pair can play the role of an intention. It (1) represents an outcome—in this case, the pleasing of James' mother—, (2) causes an event—James' going to church—; and (3) causes an event whose occurrence would normally lead to the outcome's occurrence.

It appears, then, that we can explain the relation between an action and the goal or goals to which it is directed just in terms of belief and desire. We don't need to introduce intentions as further mental states. If we like we can say that an intention just is a suitable, action-causing belief-desire pair.

This view of intention is parallel to a view about knowledge. Some suppose that knowledge is justified true belief. Or belief meeting some condition like being true and justified. On this sort of view, knowledge is not a mental state over and above belief. Rather, there are just beliefs and some of these beliefs have a special status. Similarly, Davidson's idea might be put by saying that intentions are just action-causing belief-desire pairs. (This is not to say that *any* belief-desire pair is an intention, of course.)

Now as we're about to see, this view cannot be the whole truth about intention. But at the same time I think it captures an important truth about intention, and this view, although not quite right, is a key component in a correct theory of action and intention.

5. Objection

One problem with Davidson's view—which Davidson himself raised—is that we can have intentions that don't lead immediately to action, and maybe never do. Call these cases of 'pure intending' (Davidson's term).

I might spend my whole life intending to build a squirrel house in my garden without actually doing so. There is no action-causing belief-desire pair corresponding to this intention. So the claim that all intentions are action-causing belief-desire pair must be false.

6. Intentions as all-out judgments (stored actions)

How can we respond to the objection without drastically changing the view? Here is an observation. In cases of pure intending, there is a belief–desire pair and, just like in cases of actually acting, the belief–desire pair causes and justifies the acquisition of the intention.

So we might try to solve this problem by treating intentions as something like proxies for actions. Here is the intuitive picture. The belief–desire pair causes and justifies a state which is a sort of proxy for the action intended. And then that proxy state sometimes causes the action.

This is the form of Davidson’s view.

In fact Davidson identifies the action proxy as a judgement about the desirability of a type of action.

‘the intention simply is an all-out judgement. Forming an intention, deciding, choosing, and deliberating are various modes of arriving at the judgement, but it is possible to come to have such a judgement or attitude without any of these modes applying.’
(Davidson 1980, p. 99)

We shouldn’t worry too much about this detail because there are decisive objections to this view.

7. Objection

At this point you might be wondering why we are doing this. It’s not just a history lesson. You need to understand this if you’re going to understand what intention is.

So far we have the idea of intention as a kind of proxy for action, so that an intention is what you form when you’ve decided what it would be desirable for you to do but aren’t ready to act now. The attraction of this idea is that it is relatively simple. Of course, it does involve postulating intentions as states distinct from beliefs and desires. So it isn’t as simple as the first view we considered. But, on the positive side, it hardly involves complicating the role of beliefs and desires. The only complication is that we now recognise beliefs and desires can lead not only to actions but also to intentions. But insofar as intentions are a sort of proxy for action, this is hardly a complication.

Bratman stresses this feature of Davidson’s view:

‘the basic inputs for practical reasoning about what to do—either now or later—will just be the agent’s desires and beliefs. Such reasoning, when concerned with the future, can issue in future

intentions. And these intentions are fundamentally different sorts of states from the desires and beliefs on which they are based. But there is no significant further role for these intentions to play as inputs into one's further practical thinking. Future intentions are, rather, mere spin-offs of practical reasoning concerning the future.' (Bratman 1985, p. 222)

This suggests a possible objection. Maybe there is a reason why we need to think of intentions as further inputs to practical reasoning (that is, reasoning about what to do). Maybe a model of practical reasoning which involves beliefs and desires only is insufficient. But why? How is the belief-desire model limited?

One problem with the standard belief-desire model is that it treats agents as living for the moment. There is no interesting way for the agent to act in the light of any awareness of itself as existing over a period of time during which it might perform several different actions directed to quite distinct ends.

To see why this limit matters, consider the following pair of judgements:

desire: to earn more money

belief: I can earn more money by getting a new job.

judgement: My getting a new job would be desirable.

and:

desire: to take it easy today

belief: I can take it easy today by not getting a new job today.

judgement: My not getting a new job today would be desirable.

Now there is nothing at all irrational about this combination of *judgements*. There is nothing irrational about the judgements even if we suppose that every day of your life you make the second kind of judgement. Indeed, all of these judgements may be true. This is just how it is in life. There are outcomes whose occurrence in your lifetime is desirable even though on each day of your life it would be nicer if that outcome did not occur.

But there would be something irrational about this combination of *intentions*. After all, you can know by logic alone that you will not be able to fulfil all of these intentions.

So the judgements can't be the intentions. (The combination of judgements is not irrational, but the combination of intentions is irrational.)

But Davidson's theory of intending is precisely that the judgements are the intentions. The attraction of this theory was that it involved minimal modification to the belief-desire model of action. But in the end we can't hold on to this model because it treats agents as living for the moment. And

sometimes we see our lives as involving different temporal phases and so are motivated to resolve inconsistencies in our plans. Like the inconsistency involved in intending to get a new job and, for each day, intending not to get a new job on that day.

Can we say something general about why holding this combination of intentions would be irrational? The problem is having intentions where each individual intention is perfectly rational, but the intentions taken together cannot all be fulfilled. This is captured by what Bratman calls the norm of agglomeration:

agglomeration: it is not rational to have several intentions simultaneously unless it is rational to have a single intention agglomerating them all.

To illustrate, it is not rational to intend to visit an aunt today and, simultaneously, to intend to buy shoes today unless it is rational to intend to visit an aunt today and buy shoes today.

8. Link back to decision theory

Return to the two sets of judgements. If we think of these in terms of decision theory, this is just a case where there are different ways of framing choices. We can frame the choice either in terms of whether to get a new job, or in terms of whether to get a job on each of several days. And depending on how we frame the problem, our preferences lead to different choices. And that's all that standard decision theory can say about this.

9. So what are intentions?

Elements in plans. Sui generis mental states. ***

10. What are intentions for?

'why should rational agents like us have the capacity to have *both* ordinary intentions (subject to demands for consistency and agglomeration) *and* guiding desires (which are not subject to these demands)? ... these demands [for consistency and agglomeration] are grounded largely in our needs for coordination. ... Our concern with coordination typically obliges us to form intentions, and not merely to allow desires to control our planning and conduct. This is now always the case, however: intention

formation is but one of several strategies for the resolution of practical conflicts.' (Bratman 1987, pp. 137–8)

References

- Bratman, M. (1987). *Intentions, Plans, and Practical Reasoning*. Cambridge MA: Harvard University Press.
- Bratman, M. E. (1985). Davidson's theory of intention. In B. Vermazen & M. Hintikka (Eds.), *Essays on Davidson: Actions and Events* (pp. 13–26). Oxford: Oxford University Press. Reprinted in Bratman, M. (1999) *Faces of Intention*. Cambridge: Cambridge University Press (pp. 209–224).
- Davidson, D. (1963). Actions, reasons, and causes. *The Journal of Philosophy*, 60(23), 685–700.
- Davidson, D. (1978 [1980]). Intending. In *Essays on Actions and Events* (pp. 83–102). Oxford: Oxford University Press.