Action as experience

Benj Hellie
University of Toronto

Talk at University of Miami
April 30, 2010

Overview

- I think that consciousness and agency are the same thing.
- Today we take a look at actions, and extract a number of important structural respects of resemblance with consciousness experiences (if we are willing to admit broad consciousness).
- (In the last talk Amie saw of mine, I was looking through the other end of the telescope.)

How I am thinking of action

Actions are extended: I was studying for my BA for four years.

Actions overlap: while I was studying for my BA, I did a lot of other unrelated stuff, such as spend a couple of weeks building a bicycle. We multitask.

Actions go 'live' and 'dormant'. While my building of a bicycle was live, my studies were dormant. During certain periods during the building of the bike, that went dormant and my studies became live.

We can think of actions as 'strands' coexisting within us; over time, there is variation in which strand is live.

Morals

A. Action first
B. Actions are revealed (in the good case)
C. Actions compose the subject

Plan

I. Practical and epistemic attitudes

II. The contours of action

III. Practical essential indexicality

I. Practical and epistemic attitudes

- I am going to exploit a space-age approach to epistemic modals to explain the relation between *A*-ing and trying to *A*, and then bring in intending and wanting
- Moral: action first, in the sense that successful action is (conceptually) fundamental, while trying, intending, and wanting are derivative
- In the next section we will see something stronger, namely that specific kinds of action are fundamental

Dynamic epistemic modals

Truth-conditions

$\Box P$ (I know that P) and $\Diamond P$ (so far as I know, P) perform a 'test' on some set of epistemic possibilities (mine, Fred's, yours) for a certain global condition (Veltman)

- ▶ $\Box P$: true just if at *every* possibility in the set, P;
- ▶ $\Diamond P$: true just if at *some* possibility in the set, P.

This goes naturally with the idea that it is OK for one to affirm S just when, at all one's epistemic possibilities, S

- ▶ $\Box P$ is affirmable iff:
  at all my epistemic possibilities, $\Box P$, iff:
  at all my epistemic possibilities, P:
  - ▸ After all, if the global test is passed simpliciter, it is passed at each possibility
  - ▸ So P and $\Box P$ are always 'equiaffirmable'
- ▶ $\Diamond P$ is affirmable iff:
  at all my epistemic possibilities, $\Diamond P$, iff:
  at some of my epistemic possibilities, P:
  - ▸ After all, if the global test is passed simpliciter, it is passed at each possibility
  - ▸ So whenever P is affirmable, so is $\Diamond P$, but not vice versa

Epistemic Moore's paradox

- Patterns of entailment (from the first person):
  1. P $\dashv\vdash$ $\Box P$
  2. $\neg P$ $\dashv\vdash$ $\neg \Diamond P$
  3. P $\not\vdash$ $\Diamond P$

- The paradox here is first, that (2) and (3) show a failure of contraposition; and second, that none of this goes anywhere from the third person

- The dynamic view explains all this (basically, take entailment to be 'affirmability, whenever': Yalcin)

Practical Moore's paradox

- Abbreviations:
    - A^* :=   I will have A-ed
    - $\heartsuit A$:=   I am A-ing
    - $\spadesuit A$:=   I am trying to A
- Patterns of entailment (from the first person):
    1. A^* $\dashv\vdash$ $\heartsuit A$
    2. $\neg A^*$ $\dashv\vdash$ $\neg \spadesuit A$
    3. A^* $\not\vdash$ $\spadesuit A$

Dynamic practical modals

This suggests incorporating a 'test' of one's epistemic possibilities into the semantics of practical self-attributions. Roughly:

- Taking oneself to be A-ing is taking oneself to be 'working toward' having A-ed in confidence of success:
    - $\heartsuit A$ = I am working toward A^* and $\Box A^*$
- Taking oneself to be trying to A is taking oneself to be 'working toward' having A-ed in lack of confidence of failure:
    - $\spadesuit A$ = I am working toward A^* and $\Diamond A^*$

Doing and trying

- Plausibly, then, what it is for someone to *be* A-ing at *t* is for them to be the agent of some episode of successfully having A-ed which is in progress at *t*;
  - That's not an attempt at reduction but rather a mere refraction into a different grammatical mode
- And what it is for someone to *be* trying to A (at *t*) is for them to be (at *t*) 'working toward' having A-ed.
  - In the next section I argue that 'working toward' having A-ed is derivative on successful A-ing

Practical and epistemic states

We can introduce a functor $\triangle$ which takes an attitude and makes the disposition to be in a condition indiscriminable from that attitude:

- I intend to $A = \triangle \, \heartsuit \, A$
- I want to $A = \triangle \, \spadesuit \, A$
- I believe that $P = \triangle \, \square \, P$
- It might be that $P = \triangle \, \lozenge \, P$

If this is right, there is a nice systematicity to folk psychology; which suggests that it is right.

Morals

- Intending to *A* and wanting to *A* are derivative on successful *A*-ing;
- So is trying to *A* (spot me for a moment that 'working toward' having *A*-ed is derivative);
- So *A*-ing is the fundamental practical attitude.

II. The contours of action

- What is 'working toward' having *G*-ed?
- It is that which one knows oneself to be doing whenever one takes oneself to be *G*-ing or trying to *G*;
- And this is *optimizing the excellence of one's (expected or hoped-for) completed G-ing*.
- Moral: kinds of action are conceptually irreducible and (in the good case) revealed.

Action theory old and new

- The project: explain the metaphysics of action by unpacking practical rationality
    - The old way: rationality in G-ing has something to do with optimality (Davidson)
    - The new way: rationalizations fundamentally have the form 'I'm A-ing because I'm G-ing' (Thompson)
- The old way lacks (and the new way has) a theory of what rationalizations are like, the new way lacks (and the old way has) a theory of the general structure of rationalization.
- Why can't we all just get along!

The old way
Optimizing rationalizes action

- It is rational to commence A-ing when this is the optimal live option;
- Namely, the one which maximizes expected 'value' of some universal sort:

$$EV(O) = \sum_{\pi} C(\pi/O)V(\pi)$$

  (the π are maximally specific possibilities of some sort)
- How are we supposed to operationalize this notion of universal value?

The new way
Action rationalizes action

- How does 'I'm *G*-ing' rationalize 'I'm *A*-ing'?
- Consider a conflicted person:
  - I'm buying this coach ticket to Australia because I'm tithing this year's income to the church
  - I'm buying this business-class ticket to Australia because I'm doing a triathlon there
- There seems to be no right or wrong choice for this person. Rather, the live strand selects its means in order to optimize the excellence of its eventual outcome 'as such'.

G-relative value

▶ Ordinary rationalizations of action by action seem to involve our sense of what it is for a completed action to have gone well or poorly by the standards of its kind: a good *tithing a year's income to the church* displays a certain degree of asceticism, a good *competing in a triathlon* displays a certain degree of athleticism.

▶ We want a measure of one's sense of how well a completed *G*-ing goes at a possibility by the standards of *G*-ings:

▶ $V_G(\pi)$ or '*G*-relative value' of a possibility measures this.

Optimizing an action

- 'I am A-ing because I am G-ing' is then regimented as 'what I saw in commencing A-ing when I did was that it was G-optimal', or maximized expected G-relative value out of all live options, where expected G-relative value is this:

$$EV_G(O) = C(G^*/O) \left[\sum_\pi C(\pi/G^*O)V_G(\pi) \right]$$

  - When I am confident of success, the first credence-weighting drops out;
  - That is how whether one takes oneself to be G-ing or merely takes oneself to be trying to G influences one's choice of means.

- On a strand, one is working toward having G-ed just if on that strand, one's choices are governed by their G-optimality.

First moral

- Working toward having G-ed isn't characterized in G-free terms (V_G appears in that notion but is itself inseparable from the notion of a completed G-ing);
- So no way to reduce action Davidson- or Stalnaker-style by appeal to purely nonpsychological desired outcomes;
- So kinds of actions are conceptually primitive.
- (That wraps up the argument for action-first.)

Second moral
Trying is revealed

- A revelation thesis: $\spadesuit G \equiv K \spadesuit G$
  - A self-intimation thesis: $\spadesuit G \supset K \spadesuit G$
    - Joan is trying to G. So her live stream of actions is governed by G-optimization: what she sees in each of her choices from moment to moment is that it is G-optimal. This does not seem to be compatible with Joan's reasonable suspicion that she might not be trying to G.
  - An infallibility thesis: $\spadesuit G \subset K \spadesuit G$
    - Joan is not trying to G. So her live stream of actions is not governed by G-optimization: what she sees in each of her choices from moment to moment is not that it is G-optimal. This does not seem to be compatible with Joan's reasonable suspicion that she might be trying to G.

Second moral
Action is revealed (in the good case)

- A revelation thesis: $\heartsuit G \equiv K \heartsuit G$
  - A self-intimation thesis: $\heartsuit G \supset K \heartsuit G$
    - Not plausible: unconfident trying is compatible with success
  - An infallibility thesis: $\heartsuit G \subset K \heartsuit G$
    - Not plausible: sometimes we think we will succeed but will fail (or are unreasonably confident but lucky)
- Still, in the good case one acts successfully out of expertise:
  - In expertise one is rightly confident of success;
  - When one is rightly confident of success and will succeed one knows one will succeed.
- So: **in the good case,** $\heartsuit G \equiv K \heartsuit G$.

III. Practical essential indexicality

- I am going to argue that we should sharpen the centers in centered worlds to cut within the person-slice: we want to be able to represent the perspective from within an ongoing action
- Moral: actions compose the subject, in the sense of setting ineliminable co-ordinates for the first-person perspective.
  - This is a bit more speculative than the earlier stuff . . .

Property-attitudes at subject-slices

- Content:
    - Epistemic possibilities are not possible *worlds* but possible *subject-slices*, one's 'epistemic alternatives'
    - Contents are not *propositions* but *properties* of subject-slices
- Truth:
    - Self-ascription of *F* is <u>true relative to *S*, *t*</u> iff *S* is *F* at *t*
    - *S*'s self-ascription of *F* at *t* is <u>true (simpliciter)</u> iff self-ascription of *F* is true relative to *S*, *t*

Why property-attitudes?

i. Rational distinctness despite equivalent 'objective' content
   - *glad that's over!*; *my pants are on fire*; *I'm the god on the coldest mountain*
   - Two representational states differ in rational power despite identity of how things in general might be due to differences in who one might be or when now might be

ii. Rational equivalence despite distinct 'objective' content
   - *I see a tiger*; *eat breakfast soon after waking up!*
   - The rational power of a kind of representational state might be indifferent to/'divide through by' which subject one is or time now is (considered objectively)

Property-attitudes at strand-slices

- Content:
  - Epistemic possibilities are not possible *subject*-slices but possible **strand**-slices
  - Contents are not properties *subject*-slices but properties of **strand**-slices
- Truth:
  - **Strand**-ascription of G is <u>true relative to σ, t</u> iff σ is G at t
  - σ's **strand**-ascription of G at t is <u>true (simpliciter)</u> iff **strand**-ascription of G is true relative to σ, t

Why slice the center?

Well, in general:

  i. Rational distinctness despite equivalent self-ascriptive content

     ▸ Two representational states differ in rational power despite identity of how things in general might be *and* who one might be and when now might be due to differences in which strand is live

  ii. Rational equivalence despite distinct self-ascriptive content

     ▸ The rational power of a kind of representational state might be indifferent to/'divide through by' which strand is live (considered objectively or even subjectively)

But let's see some examples . . .

Type (i): practical modes of presentation

- ▶ Familiarly, performing or simulating an action involves a different kind of understanding than reading about it, a 'practical mode of presentation'
- ▶ So two subjects could have all the same relevant objective information (and who are de se et nunc certain), who differ in that while both have some objective sense of riding a bike, only one has a live bike-riding strand
- ▶ The standard view is not completely analogous. The live strand not only serves as the circumstance of evaluation but also influences the content: plausibly, if on the live strand one is trying to chop garlic, in every epistemic possibility one is trying to chop garlic.

Type (ii): portable subroutines

- A lot of things I know how to do are 'portable' in the sense that I do them the same way within a wide range of broader practical contexts
- For example, my knowledge of how to tie a necktie is indifferent to whether my broader purpose is business or pleasure
- So the content of that knowledge 'divides through' by differences in which strand is live
- In planning my course of getting dressed, I grab the subroutine from the library and plug it in to my schedule, letting the strand take care of when the routine gets called

The logical form of action self-ascriptions

The earlier story about these lfs faces a technical problem of how to 'bind' the characterization of the action to the success ascription. Indexing to the strand allows us to bind as follows:

- ▶ $\heartsuit A$ = works toward A^* and $\square$ succeeds
- ▶ $\spadesuit A$ = works toward A^* and $\lozenge$ succeeds

Or, if the kind of strand sets the epistemic possibilities:

- ▶ $\heartsuit A$ = $\square$ succeeds
- ▶ $\spadesuit A$ = $\lozenge$ succeeds

Moral

- If we think of the subject as something like *that from which the first-person perspective emanates (over time)*;
- And if we think of **that** as *the entity relative to which our representations are most naturally assessed for truth (at a time)*;
- Then we should think of the subject as a sort of bundle of strands of action.

Conclusion

I have advanced the following morals:

A. Kinds of action are conceptually primitive;
B. The kind of the live strand is revealed in the good case;
C. Strands compose the subject.

I conclude that strands and their kinds are not easily discriminated from experiences and what they are like (in the eyes of the externalist).