

11. More on personal identity

Martín Abreu Zavaleta

July 23, 2015

We have already seen some of the puzzles and issues arising from personal identity. In particular, we have examined Locke's view quite closely so far. In his dialogue on personal identity, Perry summarizes some of the arguments and debates on personal identity that philosophers have had. Following him, we'll survey some of these views.

1 Personal identity as sameness of soul

Some people, especially in the 18th century and before, thought that there were souls: immaterial substances which could think and feel. Some people, even today, may be tempted to think that for a person at a given time to be identical to some person at another time, it is both sufficient and necessary that they have the same soul.

One initial reason why this may seem implausible is the difficulty of stating what a soul is, exactly, or how an immaterial think could causally interact with physical things, as we would expect it to do. We will examine some of those problems later in the course, when we discuss the mind-body problem. For now, we are only concerned with the issue of identity, so let's take for granted that there can be souls, and that they could interact in whatever ways they should interact with material things.

How plausible is it that identity of persons is just identity of souls? The character Gretchen in Perry's dialogue offers some arguments against this thought. Her arguments mostly have to do with issues pertaining how we know that two people are the same or not.

Her first argument is something like this, as she summarizes in p. 12:

- (1) If identity of people is identity of souls, then in order to know whether two people are identical we must know that they have the same soul.
- (2) We never know whether two people have the same soul.
- (3) So if identity of people is identity of souls, we would never know that a person at a given time is the same as some person at a different time.
- (4) But often we know when a person at a time is the same as a person at another time.
- (5) Therefore, identity of people is not identity of souls.

The character Sam tries in various ways to show how we can know that the same soul occupies a body at different times, but his attempts are not very fruitful.

One of his attempts goes as follows. Pretty much by definition, we don't have a direct means of *observing* sameness of soul. Sam claims that even though our means are not direct, we can be guided by certain psychological characteristics: memories, personality, etc. If two people have the same such psychological characteristics, then they must have the same soul.

However, Gretchen correctly points out that this is not enough evidence: how do we know that there aren't infinitely which are continuously switched so as to preserve psychological continuity? We are not in a position to know that the latter kind of situation is not the case: not by introspection, and not by indirect evidence. So it seems that the question whether two people have the same soul is simply orthogonal to the question whether they are really one and the same person.

These observations are pretty much in line with Locke's theory. In fact, Locke himself claimed that it was possible for a single person to be "distributed" over many different souls over time. He thought that the same was true of bodies.

2 Getting fancy

Sam is not a very sophisticated character in the first dialogue, but he stayed up all night thinking about a way to convince Gretchen that she can survive the death of her body. He comes out with the following objection against the thought that personal identity is just bodily identity. The objection is similar to Gretchen's objection against the "sameness of soul" hypothesis:

- (1) We can know that two things are the same person without knowing that we have the same body.
- (2) If sameness of person was just sameness of body, we would have to know that two things have the same body in order to have the same person.
- (3) So sameness of person is not sameness of body.

His argument in favor of (1) is pretty simple. When you wake up, before you open your eyes, you don't know that you have the same body that you had when you went to sleep, but you know that you are the same person. **Question:** What do you think about Sam's argument?

His second objection goes as follows resembles one of the scenarios that Locke imagines: we can imagine a situation in which a person swaps bodies. At a given time, she has a certain body, but later (perhaps as a result of some brain operation or the like) she has a different body. In fact, we may imagine, her original body could be destroyed. However, as we imagine this case, the original person wouldn't be destroyed. So sameness of person is not sameness of body.

In order for the argument to succeed, it must show not only that we can imagine the scenario that Sam presents, but that such scenario is *metaphysically possible*. Let's consider this more carefully.

Aside: varieties of possibility

We can talk about possibility in many different ways. For instance, we can talk about what is technologically possible, what is possible given our current abilities or skills, what is possible given the laws of nature, given what we believe or know, etc. Among those varieties of possibility there are three that philosophers are often concerned with: *metaphysical*, *nomological*, and *epistemic*.

Something is *nomologically possible* only if it can happen given the laws of nature and the original state of our universe. For instance, if Newton's law of inertia is in fact one of the laws that

govern the universe, then it's nomologically impossible for an object to start moving unless acted upon by an unbalanced force. However, this very same thing may be *epistemically possible* for someone who doesn't know that the law of inertia is true. Perhaps this person hasn't thought too much about the issue or he has never taken a Physics class. In any case, it is compatible with what he knows that an object may start moving without being acted upon by an unbalanced force. Then this is *epistemically possible* for him. Notice that whether something is epistemically possible depends on whose state of information we are talking about. This is not so with *nomological possibility*. Whether something is nomologically possible or not is independent of what we know or believe.

Finally, something is *metaphysically possible* if it is compatible with the most basic facts that make things what they are. For instance, it is metaphysically possible for my computer to be red, because it is compatible with the features that make it what it is, or with the nature of the computer. But perhaps it's metaphysically impossible for my computer not to be a computer, or to be a frog rather than a computer. It's usually accepted that if something is metaphysically *impossible*, then it is also *physically impossible*. However, something that is metaphysically impossible may still be epistemically possible for some people. With this in mind, let's go back to Sam's argument.

Back to the swapping case

Sam has presented us with a case in which a person swaps bodies. At the very least, his argument seems to be pretty convincing when it comes to showing that such cases are *epistemically possible*, since we seem to be able to imagine them. However, the argument requires them to be *metaphysically possible*, and it's not so clear that this follows from their epistemic possibility. Suppose that you have never heard of the celestial bodies called Hesperus and Phosphorus, and I ask you to imagine that they occupy different positions in space. However, it turns out that Hesperus and Phosphorus are one and the same planet, Venus. So the scenario I asked you to imagine is metaphysically impossible. The lesson is that we have to be careful when we talk about metaphysical possibility and arguments from what we can conceive or imagine.

After presenting these objections, Sam presents a new theory of personal identity. His proposal is somewhat more sophisticated. According to his new proposal, we should distinguish between a *person proper* and a mere *person stage*. So far we have been talking about personal identity in the following way: we consider a person at a given time, and a person at a different time, and then we ask if they are the same person.

On the new view, this is a mistake. According to that new view, a person at a given time is just a *person stage*. The actual person is something that is composed of multiple person stages. Sam uses the metaphor of a river, but we might as well use the metaphor of a road. A road is extended in space. When we point at a road, or when we are on a road, we are only in *a part* of the road. Moreover, the road is not any one of its parts. Rather, it is something like the collection of its parts, perhaps arranged in a certain way.

Sam's new view of a person draws an analogy between the person and the road. A person is not any particular one person stage. Rather, it is the whole collection of the person stages. So when we ask whether a person at a given time is the same as a person at some other time, what we are really asking is whether these two person stages are parts of the same person. What makes two stages part of the same person is that they are part of a sequence of person stages each of whom can remember enough details of some of the events experienced by her predecessor. In other words, two stages are part of the same person just in case they are psychologically continuous with each other (using the

definition of psychological continuity from last class).

However, as Gretchen points out, the new view of personal identity still has the circularity problem. So let's try to see if we can fix it by using quasi-memories. Before we had defined quasi-memories simply as states that were qualitatively identical to memories, but which didn't presuppose identity. Let's refine that. A person quasi-remembers some event E just in case:

- (a) She seems to remember E.
- (b) E in fact happened.
- (c) Her apparent memory of E is caused in the right way by traces left by some person's witnessing or performing E (not necessarily the person that quasi-remembers)

By seeming to remember I mean that the person with the quasi-memory has an experience or a mental state that is qualitatively indistinguishable from a memory: it is a sort of experience of some past event "from the inside".

This new definition still doesn't presuppose identity: two different people can quasi-remember the same thing. Notice, moreover, that every quasi-memory is a genuine memory. With the notion of quasi-memory in place, let's redefine psychological continuity using quasi-memories instead of memories:

Psychological continuity (amended): Two stages X and Y are parts of the same person if and only if they are members of the same sequence of person-stages in which each stage quasi-remembers or is able to quasi-remember enough events or enough details experienced by its predecessor.

Now that we have used the notion of a quasi-memory, the circularity objection seems to be defused. However, the view is not devoid of problems. One crucial problem, as I pointed out last class, is that we still have to determine what exactly the right way of causing a quasi-memory is supposed to be. This is not an easy task, but let's suppose there is some way to refine this. There are some more interesting problems for this view.

3 Fission cases

In the dialogue, Gretchen presents an important objection to the psychological continuity view of personal identity. She points out that, the very moment she dies, someone (e.g. some god) could create two replicas of her, each of which is psychologically continuous with her at the present time. If that were the case, which of the two replicas would be her? The psychological continuity view would say that each of them is her, since they are both psychologically continuous with her. However, the two replicas are not psychologically continuous with each other, so they are not identical to each other. We have the transitivity problem again!

Cases like the one Gretchen presents are called *fission cases* in the literature on personal identity. We could characterize a fission case as one in which:

Fission cases: At some initial time t_n , a person A is divided in two so that at time t_{n+m} , $m > 0$ there are two people (B and C) who are continuous with A, neither of which has a better claim to being A than the other.

Here are some fission cases:

- Amoeba man: Suppose that a person could split his body in two, just like amoebas do. Call this person John. At some point, John decides to split, so that there are two new bodies, each of which is physically continuous with John's and each of which is also psychologically continuous with John. Neither of the resulting amoeba men seems to have a better claim than the other to being the original John.
- Transplant: Suppose we cut your brain in half and transplant each half into a new body. The people in the new bodies each remember a lot of details about your life up to the time of the operation.
- Teletransportation: Suppose at some point in the future we manage to build a machine like the following. On one end of the machine, there is a scanner that reads all the information from your body, up to the level of subatomic structure. The scanner destroys your body as it collects the information. As the scan destroys your body, somewhere else a different machine builds a new body out of different matter from the information collected by the original scanner. One day something goes wrong, and instead of only one body being produced in the other end, two new bodies are produced in different places, each of which is an exact reproduction of the original body. These new bodies are also psychologically exactly like the original one, since their brains are qualitatively identical to the original brain.

Fission cases are very problematic, so let's examine them more carefully, focusing on teletransportation cases.

Call the person who goes into the teletransporter A and the people who come out of the teletransporters B and C, respectively. Here are the possibilities:

1. A is the same person as B, but a different person from C.
2. A is the same person as C, but a different person from A.
3. A is the same person as A, and also the same person as C.
4. A ceases to exist when she teletransported.
5. There is a composite person, C + B. A is identical to this composite person.
6. Even before the teletransportation, there were two people we called A. After the teletransportation, these two people split.

One problem for options 1 and 2 is that they seem arbitrary. Both B and C are continuous with A, so why would either of them have a better claim to being the original C? Option 3, the prediction made by the psychological continuity theory, is incompatible with the transitivity of identity. Options 5 and 6 seem a bit odd. According to option 5, after the split, A would have four legs, four arms, four eyes, etc, distributed in the bodies of B and C. According to option 6, there were two people in exactly the same places at the same time all along, which seems a bit weird.

It seems that 4 is the only option left, but even 4 seems pretty surprising. Before we started examining fission cases, the idea of teletransportation seemed pretty compelling. Now, however, the matter is not so clear. **Questions:** What if no split happened? In that case, would you survive

teletransportation? What if the teletransporter worked in a different way? What if, for instance, your body was reconstructed out of the same matter that composed it originally? what if there is a delay between the destruction of the original body and the appearance of the new bodies? Perhaps the original body is scanned first, and the new body appears somewhere else after an hour or so, at which point the original body is destroyed. Do you think these differences in the way the teletransporter works make a difference to whether one could survive teletransportation?

In the dialogue, one of the characters seems to endorse that, if there is no split, we would survive teletransportation. He rejects that psychological continuity is sufficient for personal identity, and instead he claims that in order for two people to be the same they have to be psychologically continuous with each other *and* there has to be at most one candidate for being the original person. Call this the *no competitors view*.

Gretchen presents an good objection to the no competitors view. As she points out, the no competitors view seems to make identity depend on matters extrinsic to a person: whether there are competitors or not is a matter extrinsic to me, it doesn't depend on how I am by myself. But ordinarily we would think that identity should only depend on how I am. This is a problem for the new view.

Gretchen uses her now standard argument about how we would know that we are the same person: if the no competitors view was true, then I could only know that I am the same person who was writing these notes five minutes ago if I knew that I have no competitors. But I don't know that I don't have competitors, so I don't know that I'm the same person who was writing these notes five minutes ago. This conclusion seems unacceptable.