

We've seen what Locke has to say about personal identity, admittedly only briefly, but you can see that he distinguishes between the identity of different things, in particular, distinguishing between the identity of a physical organism and a person. He's got a clear motive for doing that because of the forensic importance of the notion of person, its relation to morality and desert, and so forth. He links personhood with consciousness and therefore with memory.

Now, Thomas Reid famously raised a problem case for Locke, which he put like this: suppose you've got a young lieutenant in the Army. He can remember what he did as a child, maybe, for example, he can remember stealing an apple, maybe feels guilty about it. When he becomes a general, the general can remember what the lieutenant did. But he can't remember what the child did. He's completely forgotten about the Apple incident. Now, if memory constitutes personal identity, then it looks like the lieutenant is the same one, the same person, as the child. The general is one and the same person as the lieutenant. But the general is not the same person as the child. And that seems to be a contradiction because the notion of identity is supposed to be a transitive relation. That is, if X is identical with Y and Y is identical with Zed, it follows that X is identical with Zed. That seems to be a logical truth about the notion of identity. Indeed, it follows from Leibniz's Law.

Well, there is a nice, elegant way of getting around this. So we introduce what's called the ancestral relation. You'll see it's obvious why it's called the ancestral relation because the classic example of this is the relation of being an ancestor. So look at this definition: "X is an ancestor of Y if either X is a parent of Y or X is a parent of an ancestor of Y". Now, when you first see that definition, you might think, isn't that a bit odd? Because the word ancestor is not only appearing in the word to be defined, it's also appearing in the definition. Isn't that circular? Well, in a sense, it is, but it's not viciously circular. And that's because we have the first clause as well. So suppose we want to identify who my ancestors are. Well, what this says is, first of all, my parents are ancestors, okay? But then their parents are ancestors because they are then parents of ancestors of mine. And then their parents are ancestors of mine because they're parents of ancestors of mine, and so on. So actually, it's a very elegant and neat way of getting a recursive definition. The definition of ancestor appeals to the notion of ancestor itself, so it's recursive, but it's not viciously so.

Now consider that in the abstract. Consider what we've done. We've taken the notion of parent, the relation of being a parent, and we have generated from that a more general relation that includes parenthood, but as it were, iterates it. So we call ancestor the ancestral relation of parent.

Now suppose we do the same with memory. Suppose we say that X is memory continuous with Y if either X can remember Y's actions or X can remember the actions of someone who is memory continuous with Y. So now what we have is that the general is memory continuous with the child because the general is memory continuous with the lieutenant, and the lieutenant is memory continuous with the child. So we get around Reid's problem by extending Locke's account. We don't make it rely just on what a single-step memory. We allow iterated memories to count as well.

There's another advantage of this sort of approach, which is as follows: my continuity with myself yesterday. We're reluctant to think that that's anything to do with whether I can remember exactly what I did yesterday. I mean, suppose you go to a really good party and afterwards you can remember very little of what went on. That doesn't mean you're not the same person. Now, suppose, on the other hand, you imagine yourself throughout those hours. If your consciousness is continuous throughout the period, then you can, at each moment, remember what happened in the immediately previous moment. You have this unbroken continuity of experience, and the ancestral relation lets you build that right through the day. So you don't, at any point, have to say that, at 11 o'clock in the evening, you have to be able to remember what happened at 9 o'clock in the morning. The continuity can come through the ancestral relation.

But there are, of course, lots of problems remaining. One problem is so-called quasi-memory. If you base personal identity on memory, the trouble with that is something only counts as a genuine memory if it concerns your own experiences. Suppose I wake up with an apparent memory of your experiences. Very peculiar. But would we call that memory? No, we wouldn't. We'd say that isn't a genuine memory. Something's gone wrong, goodness knows what, but that's not a genuine memory. Okay, so straight off we need to say that what we're using as a criterion of personal identity is not genuine memory. It's apparent memory, quasi-memory. If we say genuine memory, then that's just going to be circular because we're only going to count it as a genuine memory if we have personal identity.

Another problem is sleep. Well, continuity of consciousness may do the job from morning until night, but what about when you go to sleep? Does your consciousness then switch off? Well, maybe you can get around that. Maybe you can say, with Descartes, that it's of the essence of your mind that it's always thinking. Maybe sometimes thinking semi-consciously. And maybe you're going to be able to build that into some account. You can see there's a difficulty there. We're not going to be able to rely on full consciousness in the way that Locke wanted to.

But what about coma? What if someone goes into a coma? Suppose for a time, brain activity ceases. We have every reason to think that there is no consciousness whatsoever over that period. But then it revives. Would we not want to say that it's the same person, particularly if they continued to act in the same way, they retain memories from before, and so on? Well, this suggests that some element of bodily or at least brain continuity is desirable. It enables you to bridge over the in these sorts of cases. If somebody had a really bad accident, the brain appeared to switch off for a time but then switched on again, and things continued more or less as normal, I'm sure we would want to say it's the same person. We do that because it's largely because it's the same physical body, and we assume that with the same physical body, broadly the same mind is there. It's like turning a computer off and turning it on again. It's still your computer. It still has all the memories and software, etc., from before. And that might lead us to think that bodily identity is really the crucial thing. Consciousness doesn't matter so much. What matters is the physical seat of consciousness, the thing that actually causes it.