



Self-Perpetuating Process as a Resolution of Certain Identity Paradoxes

Carl M. Johnson

Logic

May 2007

Contents

I. Introduction	2
II. The problem with using formal identity	3
III. The problem with pure nominalism	5
IV. The problem with simple co-material identity	6
V. Continuous recognizability	7
A. Organ donation example	7
B. The problem of splitting	9
VI. Self-perpetuating identity	10
VII. Application to paradoxes	13

I. Introduction

In the last footnote of his chapter of *Paradoxes* on vague objects, R. M. Sainsbury presents a modified version of the sorites paradox directed at the human being: A human being still exists even if he or she lose one molecule from his or her body. Losing one molecule cannot make the difference between existence and non-existence for a human being. Therefore, even if all your molecules are taken away one at a time, you will still exist even if none of your molecules do. However, it is absurd to suppose there could be a human being without a body; therefore, human beings do not exist.

In this same section, Sainsbury presents the familiar example of the ship of Theseus which is slowly replaced plank by plank. The reader is asked, “Did Theseus’ ship survive?” If one had made a new ship of the discarded planks, “Does this have a better claim to be the original ship of Theseus?” asks Sainsbury.

In this paper I will show that both paradoxes can be partially defused by creating a more precise definitions of “human being,” “ship,” and other relevantly similar composite objects that are based on continuous self-perpetuating processes rather than static substances. The use of a more precise definition for these types of objects does not rely on process ontology being the correct underlying metaphysical reality and is quite independent of whatever the particulars that underlie our world turn out to be. Once this more precise definition is in place, paradoxes of the sort mentioned above can be safely sidestepped by denying their assumptions.

II. The problem with using formal identity

Underlying the two paradoxes presented so far is the central question, How can there be things which are not the same as their material composition yet which nevertheless rely on their material composition for their existence? Another pertinent example in this regard is the river. Since the time of Heraclitus, philosophers have been familiar with the problem with supposing the material identity of rivers. A river cannot be just what makes it up at a particular time, since if we say that a river is the matter that makes it up, then we must agree with Heraclitus that one cannot step in the same river twice, because a river's material composition is in constant flux. At the same time, we commonly speak of rivers as though they were things that persist for years and years.

To begin, we should note that "same" by itself is an ambiguous term. In our most typical usage of "same," we only mean "same in certain respects." For to say that a river at one time is not the *exact* same as the river at another time is tautological. Obviously, the latter river has at least one a different property than the first river, since it has the property of being at the later time, where the first river has the property of being at an earlier time. Even if one were to discount being at a particular time as a relevantly meaningful property for whatever reason, there would still be that issue that if anything else in the universe has moved between the two times, then the first and second rivers are different with respect to their distance from the objects in the universe that have moved. Of course, such properties as the distance between one river and everything else in the universe are clearly not essential properties for the identity of a river, but the question at hand is exactly what properties are essential for sameness in our common usage of the term, so we cannot dismiss this problem out of hand. Thus, the "sameness" of being the same river (or the same anything else) at two different times must not mean "same in all respects" but only "same in certain relevant respects."

One way to resolve this ambiguity and solve the problem brought up by Heraclitus is to say that ordinary references to sameness in rivers refers to sameness with respect to

one of the river's formal physical properties even if its material make up varies. However this quickly runs into difficulties. Imagine that there existed perfect teleporting machines that could duplicate objects down to their most minute physical properties and that we used such a machine to make a perfect physical duplicate of the Colorado River on Mars. Each water molecule in the teleported duplicate would be positioned as the exact analog of our terrestrial river, and the banks as well would be perfectly copied. Nevertheless, in spite of being in every way like the Colorado River, our duplicate river would not be the Colorado River for the simple fact that it is located on Mars and not on Earth, which is commonsensically essential to the identity of the Colorado River.

Accordingly, we might next propose that geography or physical location is essential to the identity of the river. Unfortunately, this will just as quickly run into problems, since a particular river's banks are constantly changing and being eroded. For our example of the Colorado River, at one time the river's banks were on the surface of a plane but now they are the bottom of a canyon. Thus, the formal characteristic shared by a river over time cannot be just its banks, since they are not the same over time. Nor can it be the generalized path of the river, since this is also subject to radical change, as when a river breaks through and finds a new course or when a river alternates between meandering and bypassing oxbow lakes.

One might protest that membership in certain classes allows some flexibility. For example, something can be considered "triangular" for everyday purposes even if it is not precisely composed of three straight lines, but instead it is composed of atoms lined up with a certain degree of tolerance. The river problem though does not allow recourse to geographic formalism by means of increased tolerances. Consider a case in which two rivers happen to run parallel to one another from source to the sea while separated by only a few meters. In this case, we would say there are two rivers, not one. At the same time, if one river were to shift over by a few meters such that its new course is parallel to its old course, it would still be one river, not two. Thus, merely adding tolerances is

not sufficient to resolve the problems with regarding geography as the essential formal property of rivers. Tolerances are allowable as means of defining classes of objects but insufficient when distinguishing between individual objects.

A river also cannot even be any such continuous stream of water that shares a common origin, since it is not uncommon for one source of a river to dry up, but for the river to continue by using different sources. So long as the main stream continues to flow, a river can still be the same river even if fed by different waters.

Thus, none of the formal properties listed so far—particle properties, geographic location, or head water source—can fully account for our understanding of rivers.

III. The problem with pure nominalism

One response to these many difficulties is to abandon all hope of coming up with a precise definition of river or ship or human being and switching to a form of pure nominalism about the matter. In this case, one says that a river is whatever people say it is and there's nothing more to it than that. If we cannot come up with a precise definition of river, then the problem is with the natural imprecision of human languages and nothing can be done to fix it.

This pessimistic account does contain some correct observations, but on the whole it sacrifices too much. Yes, it is difficult to come up with a precise definition of river, but nevertheless it is necessary. Suppose a river runs through two states, and they form a contract stating that each is allowed to use a certain amount of water from the river. Then suppose that after a flood causes a shift in the course of the river, one state suddenly begins to take more than its share of water but argues, "Who is to say we are in the wrong? The contract we formed was about a certain river, but that river is gone now. The contract can no longer bind us since its object does not exist." Certainly, we would like to have recourse to a precise means by which to settle such a legal dispute. On what grounds can the wronged state claim that the river still exists?

However, even this does not capture the full problem of pure nominalism. Of course words naturally mean just what we as a society want them to mean. Thus the nominalist is correct that it may be the case that our use of some terms is so jumbled that their contradictions and ambiguities cannot be removed. Accordingly, if we wanted to define a river as being the same only if it had the same atoms or the same particle properties or the same banks or same general location or whatever else, we would be entitled to use that definition if we wanted. The goal here, however, is to produce a definition that allows us to preserve our commonsense understanding of terms while still being able to reason rigorously using these terms. The river example is just a test case for a broader class of similar objects that exist as sets without fixed membership. If we cannot define river, then it is likely that we will be even harder pressed to define a number of other terms. The end result will be the inability to make more than suggestive claims about individual objects in the world. For this reason, it is worth making every effort at capturing an appropriate definition for a term before resigning ourselves to linguistic chaos.

IV. The problem with simple co-material identity

One way to solve the river problem might be to propose a theory of co-material identity. In this theory, as long as a certain arbitrarily defined percentage of material from the original composite object at time 1 remains in the object at time 2 then the reconstituted object is the same. Thus for the ship of Theseus example, the modified ship is the same ship so long as at least a particular number of original planks remain. It will then be up to the users of the word "same" to determine whether for their purposes "same" requires just one original plank or at least half of the original ship's composition be unchanged. The difference is arbitrary and can be disambiguated as needed, eg. for legal purposes connected to the ship's docking rights one plank might be sufficient, but for the purposes

of a museum the “same” ship should be composed of at least three quarters original material.

There are two obvious problems with co-material identity. First, it is not clear whether a ship constructed out of discarded planks would have the same right to the name “ship of Theseus” as the modified ship. It is possible that our definition of sameness will sometimes allow splitting (see below), but we want to avoid this possibility where we can. Another difficulty would come if the modified ship and discarded parts ship both contained fifty percent of the original material. It might seem in this case that which of the two is called the ship of Theseus can be changed by taking a single plank from one ship and putting it on the other and thus changing the balance of parts. However, the exchange of one plank seems like too small of a change for identity to be passed with it.

If the ship of Theseus example is a difficult one for the co-material theory of identity, then the river example is deadly. It should be clear enough that the Colorado River is rapidly emptying into the ocean. Whenever it would be that we designate the reference set of material in the Colorado River by which to make later comparisons, after a certain amount of time, more of the original river would be in the ocean than contained in what we take to be its banks, and after a little more time none of the original river would be left. Thus simple co-materiality is not enough to give a rigorous definition of commonsensical identity.

V. Continuous recognizability

A. Organ donation example

Before proposing a new model to solve the various problems listed so far, it might be useful to give some examples that demonstrate what our common sense requirements for identity are.

Consider for the sake of example a case in which one person, Alice, donates her kidney to her friend, Berta. After the operation, we will naturally want to say that legally

the kidney is now Berta's even though historically we can recognize that it was once Alice's. That is we can say that Berta's new kidney is the same as Alice's old kidney. (Also we will say Berta is still Berta in spite of her now containing a new organ. Thus, she is, in one sense of the term, the same person after the operation.) Furthermore the one kidney will continue to have the same two properties (legally being Berta's and historically being Alice's) even if after years of being in Berta's body all of the atoms within the kidney just happen to be replaced by new atoms from the food that Berta eats. Much as Berta retained her identity even with a new kidney, the kidney retains its identity even after its molecular parts are replaced. Thus, we can say that neither property is strictly based on material identity. Perhaps we might then propose that the identity of the kidney is based on one of its particular formal properties—in this case that it has Alice's DNA patterns even if the atoms making up that pattern are different.

However, as was the case with the river example, formal properties are not a reliable carrier of identity. Suppose that Alice and Berta had been identical twins with identical genomes. In this case the formal property that we are using to distinguish the kidney as being historically Alice's is no longer sufficient to distinguish parts of Berta's body that have the property of being historically hers from those that do not. Thus, if we wanted to say that a person is the same over time because they have the same DNA patterns, we would be unable to distinguish between twins and would have to hold that cells with DNA mutations are not a part of the same person.

To give a slightly macabre example of the difficulties of formal properties, suppose after removing Charlie's kidney instead of just implanting the organ in Derek mad scientists feed Charlie's kidney to Derek and these atoms from Charlie's kidney just so happened to replace all the atoms in Derek's kidney by moving to places in Derek's body analogous to their former place in Charlie's. Unlike the case of Alice and Berta, we would not say that Derek's new kidney is the same as Charlie's old kidney even if it takes on a shape similar to its old shape, because we would hold that at the time of digestion

Charlie's kidney was destroyed and any new kidneys that come about cannot be inherit the sameness of the destroyed kidney. We would only say that the atoms in Derek's new kidney are the same as the atoms in Charlie's old kidney. Thus, the sameness of a kidney over time is not merely its physical shape just as it was not its DNA.

The difference between Berta's new kidney and Derek's new kidney is illustrative of the operant properties that underlie common sense identity. Berta's kidney is the kidney that was Alice's in part just because if we could have theoretically watched the kidney continuously existing as the same kidney even after it was moved into place, whereas Charlie's kidney is visually destroyed during digestion. So at least partially, our common sense notions of identity fall out of the particular visual object recognition heuristics that are present in the human brain. Accordingly, one model of sameness that we might employ is continuous recognizability. If we could (at least in theory) watch something throughout its transformation and recognize it the whole time then it is the same as the original thing. The problem with continuous recognizability as a theory of identity is that it is just a reification of our prejudices. That is, if the question is, "Is this the same object as the one that was here before?" the only advice that continuous recognizability gives is, "Yes, if you continuously thought so. No if not." Thus, it is only useful in the cases where we can already recognize sameness, but it does not tell us how the recognition process actually works. Recognition itself is a black box function, and identity is just the sameness of the function's output over time.

B. The problem of splitting

The difficult with leaving recognition as a black box is that there are cases where our intuitions seem to clash, and continuous recognizability does little to resolve the problem. Consider Aristotle's acorn and oak example. We will say that the acorn is the same as the oak, in part because we could theoretically watch the acorn become the oak, as is done in time lapse photography videos. If we watch further, we might see

the oak cut down and turned into a plank of lumber. However, we might be reluctant to say that an oak is the “same” as a plank in a manner analogous to the way that an oak is the “same” as its originating acorn. The sameness of the oak to the plank is purely a material sameness, but the sameness of the oak to the acorn is a different sort of sameness, such that transitivity does not hold between the acorn and the plank. Thus, although continuous recognizability holds from the acorn to the oak to the plank, “sameness” can be broken into two different sorts, one type of which holds from the acorn to the oak, the other type of which holds from the oak to the plank. Without understanding what it is that we are recognizing, it will not be possible to differentiate these two kinds of sameness.

Similarly, consider a variation on the river example. The river that is the Missouri River in Montana is the same river as the Mississippi River in Louisiana, and the river that is the Ohio River in Pennsylvania is the same as the Mississippi River in Louisiana, but the Missouri River is not the same as the Ohio River in certain relevant respects. If possible, we want our definition to allow for the non-transitive sameness of rivers and tributaries, but doing so on the basis of just continuous recognizability is impossible, since someone walking down from the Missouri River’s banks along the Mississippi and on to the Ohio River will recognize what is being seen as the same river throughout the journey. The changes that make the difference between the Missouri and the Ohio will not be picked out by continuous recognizability alone.

VI. Self-perpetuating identity

One way to fill in the recognition gap is to say that what is recognized is the self-perpetuation of a particular pattern. Consider the river example. If we define a river as a process by which the gravitationally influenced motion of certain water molecules causes other water molecules to follow them in a similar manner then we can solve many of the problems concerning the sameness of rivers. If we have a river where molecule A is replaced

by molecule B, we can recognize the two of them as participating in the same river if we see that it was the gravitationally compelled motion of molecule A that caused the motion of B, which in turn will allow the motion of its successor C. This sameness will hold even if the river changes its shape, so long as the changes do not interfere with the process by which the void left by water molecules moving due to gravity is filled by other molecules of water that are also gravitationally compelled. Similarly, the river's source changing is fine as long as the water that enters the river from the new source is similarly compelled to run downhill by the opening of space left by the vacancy of its predecessors. The identity of the river does rely partially on co-materiality, but only from moment to moment within the self-perpetuation of the process. Even if all of the river's present molecules are someday gone the river will remain if there is an unbroken chain of causality connecting the current molecules to the molecules that will someday inherit the title of "the same river."

Next consider the ship of Theseus. In this case, we can define a "ship" as the material that inherits the process of being treated as a seagoing vessel by a crew or holding a certain form. Note that the ship need not always be at sea or have a crew, just that it must be either treated as a ship by some crew or be a collection of materials that base their shape on the prior shape of the ship. That is, as long as the ship is at sea and being used by its crew, it is the same vessel even if the crew replaces some of its parts, so long as they continue to use the new set of parts as a ship, and when it is uncrewed, it is fine for some of its material to be replaced by new material (eg. through planks rotting), so long as the new material bases its shape on the old within a certain degree of tolerance. Thus, so long as the process by which the parts of the ship are replaced is slow and continuously a part of the process of using the ship as a ship or the ship be weathered by the elements, it is the same ship. Hence taking a plank off the back of a ship and nailing it to a different ship does not make the other ship "the same ship" as the first, since the plank was not being used as a ship during the time between its removal from the first

ship and its addition to the second. Similarly, constructing a ship out of junkyard scraps is not sufficient to inherit the title of the same ship.

One consequence of the proposed definition of identity in ships is that if a ship is dry docked, taken completely apart, and left for scrap, but then reassembled using all of and only its original parts by an unaffiliated crew, then it is no longer the same ship, whereas it may be necessary in the normal service life of a ship to take it apart and reassemble it in a similar manner, but because of the continuity of treatment by its crew, this second ship will be considered the same despite the superficial similarity of its treatment. This consequence holds whenever the ship is dismantled to the point that it is no longer called a “ship” rather “a collection of ship parts.” It may seem counterintuitive, but I hold that this result is actually a useful feature of my proposed sharpening of the definition of identity for certain composite objects. While it may be controversial, common sense will not totally reject our saying that a newly constructed ship is a new ship, even if old parts were used. Part of the continuity of sameness in retrofitted ships is that they have a continuous treatment as a ship by some community. Lacking that treatment, the identity of the ship parts as parts of a particular ship, rather than parts in general, evaporates. That the sharpening of our definition lets us make a meaningful decision about this potentially controversial choice is a benefit of our proposed definition, not a shortcoming.

Consider again the example of the kidneys. Our proposal is that a kidney is the same kidney so long as the biological processes that cause an organism to sustain its existence are still in effect. As with the reassembled ship, we hold that kidney that just so happens to be reconstituted is not the same kidney. Here our intuition more closely aligns with the result of the proposed definition, giving more confidence to the reassembled ship example. Similarly, we allow that the kidney is the same kidney that used to be in someone else even now that it is implanted in its new host. Similarly, the patients can be considered the same people after their surgeries, even if they receive new organs

with different atoms, different DNA, etc., so long as their life processes as human beings continue without interruption.

With the example of the oak, we may now distinguish clearly between the biological sameness of the acorn and the tree and the material sameness of the oak and the plank by defining biological sameness according to the self-perpetuating properties of cells, and material sameness as the self-perpetuation of atoms. (Note that since quantum mechanics tells us that atoms contain numerous “virtual particles” that pop in and out of existence ever second and help give the nucleus its shape, it is not practical for us to require the material sameness of the plank to hold down to the sub-atomic level. Instead it must merely hold to it down to the atomic level.)

Finally, we can at last pick out the difference between the Missouri River and the Ohio River. When the water in the lower Mississippi River flows out, it leaves a void that water from both of the rivers rushes in to fill. However, there is no unidirectional causal chain between the Missouri River and the Ohio. Water evacuating from the one river in no way contributes to the evacuation of water in the other river. Thus, we can distinguish between the two rivers, even though they become a part of the same river further downstream.

VII. Application to paradoxes

Returning to the paradox that opened the paper, we said that, “A human being still exists even if they lose one molecule from their body.” However, in light of our newly refined definition, we cannot assent to that assumption as it is stated. It may sometimes be the case that a person persists even after the loss of a molecule, but if this is the case, it is only the case in virtue of the fact that this molecular loss had not disrupted the larger self-sustaining processes within the human being. Instead we must say, “A human being still exists even if they lose one molecule from their body, *so long as the loss of this molecule does not interrupt the biological self-perpetuation of the person as an organism.*” With that clause

added to the assertion, the paradox collapses, since it no longer allows itself indefinitely repeated application but only strictly limited application, such that the rate of molecular loss is never significantly greater than the rate of molecular replacement.

Similarly, we can answer the question, “Does [the ship built from discarded parts] have a better claim to be the original ship of Theseus?” with a resounding “no,” given the stipulation that when the ambiguous term “original ship” is used it refers to sameness according to our criterion described above rather than sameness according to a naïve materialistic criterion. Hence this paradox turns out to be a matter of the ambiguity of the term “same” rather than a matter of vagueness within the term.

With these paradoxes safely sidestepped, we can avoid the recourse to completely subjective nominalism and continue to use our commonsensical terms with confidence that if necessary there is recourse to more precise formulations of the same basic understanding of the world.