

Identity Through Possible Worlds: Some Questions

RODERICK M. CHISHOLM
BROWN UNIVERSITY

It is now easy to see a simple way of avoiding undesirable existential generalizations in epistemic contexts. Existential generalization with respect to a term—say *b*—is admissible in such contexts if *b* refers to one and the same man in all the “possible worlds” we have to consider.¹

In an article on Hintikka’s *Knowledge and Belief*, I suggested that certain difficult questions come to mind when we consider the thought that an individual in one possible world might be identical with an individual in another possible world.² The present paper is written in response to the editor’s invitation to be more explicit about these questions.

Let us suppose, then, that the figure of an infinity of possible worlds makes good sense and let us also suppose, for simplicity of presentation, that we have a complete description of this one. We may consider some one of the entities of this world, alter its description slightly, adjust the descriptions of the other entities in the world to fit this alteration, and then ask ourselves whether the entity in the possible world that we thus arrive at is identical with the entity we started with in this world. We start with Adam, say; we alter his description slightly and allow him to live for 931 years instead of for only 930; we then accommodate our

¹ Jaakko Hintikka, *Knowledge and Belief: An Introduction to the Logic of the Two Notions* (Ithaca: Cornell University Press, 1962), p. 152.

² “The Logic of Knowing,” *Journal of Philosophy*, LX (1963), pp. 773–795; see especially pp. 787–795.

descriptions of the other entities of the world to fit this possibility (Eve, for example, will now have the property of being married to a man who lives for 931 years instead of that of being married to a man who lives for only 930); and we thus arrive at a description of another possible world.³

Let us call our present world “W¹” and the possible world we have just indicated “W²”. Is the Adam of our world W¹ the same person as the Adam of the possible world W²? In other words, is Adam such that he lives for just 930 years in W¹ and for 931 in W²? And how are we to decide?

One’s first thought might be that the proposition that Adam is in both worlds is incompatible with the principle of the indiscernibility of identicals. How could our Adam be identical with that one if ours lives for just 930 years and that one for 931? Possibly this question could be answered in the following way:

“Compare the question: How can Adam at the age of 930 be the same person as the man who ate the forbidden fruit, if the former is old and the latter is young? Here the proper reply would be: it is not true that the old Adam has properties that render him discernible from the young Adam; the truth is, rather, that Adam has the property of being young when he eats the forbidden fruit and the property of being old in the year 930, and that these properties, though different, are not incompatible. And so, too, for the different possible worlds: It is not true that the Adam of W¹ has properties that render him discernible from the Adam of W²; the truth is, rather, that Adam has the property of living for 930 years in W¹ and the property of living for 931 in W², and that these properties, though different, are not incompatible.”

I think it is clear that we must deal with the old Adam and the young Adam in the manner indicated; but in this case, one could argue, we know independently that the same Adam is involved throughout. But are we justified in dealing in a similar way with the Adam of W¹ and the Adam of W²? In this latter case, one might say, we do not know independently that the

³ It should be noted that the possible world in question is not one that Hintikka would call *epistemically* possible, for it could be said to contain certain states of affairs (Adam living for 931 years) which are incompatible with what we know to hold of this world; hence it is not one of the worlds Hintikka is concerned with in the passage quoted above. But it is *logically* possible, and that is all that matters for purposes of the present discussion.

same Adam is involved throughout. Here, then, is one of the questions that I do not know how to answer. Let us suppose, however, that we answer it affirmatively.

The Adam of this world, we are assuming, is identical with the Adam of that one. In other words, Adam is such that he lives for only 930 years in W^1 and for 931 in W^2 . Let us now suppose further that we have arrived at our conception of W^2 , not only by introducing alterations in our description of the Adam of W^1 , but also by introducing alterations in our description of the Noah of W^1 . We say; "Suppose Adam had lived for 931 years instead of 930 and suppose Noah had lived for 949 years instead of 950." We then arrive at our description of W^2 by accommodating our descriptions of the other entities of W^1 in such a way that these entities will be capable of inhabiting the same possible world as the revised Noah and the revised Adam. Both Noah and Adam, then, may be found in W^2 as well as in W^1 .

Now let us move from W^2 to still another possible world W^3 . Once again, we will start by introducing alterations in Adam and Noah and then accommodate the rest of the world to what we have done. In W^3 Adam lives for 932 years and Noah for 948. Then moving from one possible world to another, but keeping our fingers, so to speak, on the same two entities, we arrive at a world in which Noah lives for 930 years and Adam for 950. In that world, therefore, Noah has the age that Adam has in this one, and Adam has the age that Noah has in this one; the Adam and Noah that we started with might thus be said to have exchanged their ages. Now let us continue on to still other possible worlds and allow them to exchange still other properties. We will imagine a possible world in which they have exchanged the first letters of their names, then one in which they have exchanged the second, then one in which they have exchanged the fourth, with the result that Adam in this new possible world will be called "Noah" and Noah "Adam." Proceeding in this way, we arrive finally at a possible world W^n which would seem to be exactly like our present world W^1 , except for the fact that the Adam of W^n may be traced back to the Noah of W^1 and the Noah of W^n may be traced back to the Adam of W^1 .

Should we say of the Adam of W^n that he is identical with the Noah of W^1 and should we say of the Noah of W^n that he is identical the Adam of W^1 ? In other words, is there an x such that x is Adam in W^1 and x is Noah in W^n , and is there a y such

that y is Noah in W^1 and y is Adam in W^n ? And how are we to decide?

But let us suppose that somehow we have arrived at an affirmative answer. Now we must ask ourselves: How is one to tell the difference between the two worlds W^1 and W^n ? Shall we say that, though they are diverse, they are yet indiscernible from each other—or, at any rate, that the Adam of W^1 is indiscernible from the Adam of W^n (who is in fact the Noah of W^1) and that the Noah of W^1 is indiscernible from the Noah of W^n (who is in fact the Adam of W^1)? There is a certain ambiguity in “discernible” and in “indiscernible”. The two Adams could be called “discernible” in that the one has the property of being Noah in the other world and the other does not, and similarly for the two Noachs. But in the sense of “indiscernible” that allows us to say that “Indiscernibles are identical” tells us more than merely “Identicals are identical,” aren’t the two Adams, the two Noachs, and the two worlds indiscernible? Could God possibly have had a sufficient reason for creating W^1 instead of W^n ?

If W^1 and W^n are two different possible worlds, then, of course, there are indefinitely many others, equally difficult to distinguish from each other and from W^1 and W^n . For what we have done to Adam and Noah, we can do to any other pair of entities. Therefore among the possible worlds which would seem to be indiscernible from this one, there are those in which you play the role that I play in this one and in which I play the role that you play in this one.⁴ (If this is true, there may be good ground for the existentialist’s *Angst*; since, it would seem, God could have had no sufficient reason for choosing the world in which you play your present role instead of one in which you play mine.)

Is there really a good reason for saying that this Adam and Noah are identical, respectively, with that Noah and Adam? We opened the door to this conclusion by assuming that Adam could be found in more than one possible world—by assuming that there is an x such that x is Adam in W^1 and lives here for 930 years and x is also Adam in W^2 and lives there for 931. If it is reason-

⁴ “She (Ivich) looked at the glass, and Mathieu looked at her. A violent and undefined desire had taken possession of him; a desire to *be* for one instant that consciousness . . . to feel those long slender arms from within. . . . To be Ivich and not to cease to be himself.” Sartre, *The Age of Reason*. Compare N. L. Wilson, “Substance without Substrata,” *Review of Metaphysics*, XII (1959), and A. N. Prior, “Identifiable Individuals,” *Review of Metaphysics*, XIII (1960).

able to assume that Adam retains his identity through the relatively slight changes involved in the transition from W^1 to W^2 , and so, too, for Noah, then it would also seem reasonable to assume that each retains his identity through the equally slight changes involved in all the other transitions that took us finally to W^n . (These transitions, of course, may be as gradual as one pleases. Instead of it being a year that we take away from Noah in our first step and give to Adam, it could be only a day, or a fraction of a second.) But identity is transitive. And therefore, one might argue, once we allow Adam to exist in more than one possible world, we commit ourselves to affirmative answers to the puzzling questions we have encountered.

Is there a way, then, in which we might reasonably countenance identity through possible worlds and yet avoid such extreme conclusions? The only way, so far as I can see, is to appeal to some version of the doctrine that individual things have essential properties. One possibility would be this:

For every entity x , there are certain properties N and certain properties E such that: x has N in some possible worlds and x has non- N in others; but x has E in every possible world in which x exists; and, moreover, for every y , if y has E in any possible world, then y is identical with x . (If "being identical with x " refers to a property of x , then we should add that E includes certain properties other than that of being identical with x .) The properties E will thus be *essential* to x and the properties N *non-essential*, or accidental.⁵

To avoid misunderstanding, we should contrast this present use of "essential property" with two others.

(1) Sometimes the "essential properties" of a thing are said to be just those properties that the thing has *necessarily*. But it is not implausible to say that there are certain properties which are such that *everything* has those properties necessarily; the properties, for example, of being either red or non-red, of being colored if red, and of being self-identical.⁶ Thus the Eiffel Tower

⁵ We could put the doctrine more cautiously by saying that the distinction between the two types of property holds, not for *every* entity x , but only for *some* entities x . But what reason could there be for thinking that it holds of some entities and not of others?

⁶ Sometimes these properties are called "analytic properties" or "tautological properties"; but the property of being colored if red should not be so-called if, as some have argued, "Everything that is red is colored" is not analytic.

is necessarily red or non-red, necessarily colored if red, and necessarily self-identical; and so is everything else.⁷

(2) And sometimes it is said (most unfortunately, it seems to me) that each individual thing is such that it has certain properties which are essential or necessary to it “under certain descriptions of it” and which are not essential or necessary to it “under certain other descriptions of it.” Thus “under one of his descriptions,” the property of being President is said to be essential to Mr. Johnson whereas “under that description” the property of being the husband of Lady Bird is not; and “under another one of his descriptions,” it is the other way around. Presumably *every* property P of every individual thing x is such that, “under some description of x,” P is essential or necessary to x.

But if E is the set of properties that are essential to a given thing x, in the sense of “essential” that we have defined above, then: E will not be a universal property (indeed, *nothing* but x will have E); some of the properties of x will not be included in E; and E will not be such that there are descriptions of x “under which” E is not, in the sense defined, essential to x.

If we accept this doctrine of essential properties, we may say, perhaps, that the property of living for just 930 years is essential to Adam and therefore that he may inhabit other possible worlds without living for just 930 years in each of them. And so, too, perhaps, for having a name which, in English, ends with the letter “m”. But, we may then go on to say, somewhere in the journey from W^1 to W^n , we left the essential properties of Adam (and therefore Adam himself) behind. But where? What *are* the properties that are essential to Adam? Being the first man? Having a name which, in English, begins with the first letter of the alphabet? But why *these* properties? If we can contemplate Adam with slightly different properties in another possible world, why can’t we think of him as having ancestors in some possible worlds and as having a different name in others? And similarly for any other property that might be proposed as being thus essential to Adam.

⁷ From the proposition that the Eiffel Tower is red and necessarily colored if red, it would be fallacious to infer that the Eiffel Tower is necessarily colored; this is the fallacy of inferring *necessitate consequentis* from *necessitate consequentiae*. And from the proposition that the Eiffel Tower is necessarily red or non-red, it would be fallacious to infer that the proposition that the Eiffel Tower is red or non-red is a necessary proposition; the proposition could hardly be necessary, for it implies the contingent proposition that there is an Eiffel Tower. This latter fallacy might be called the fallacy of inferring *necessitate de dicto* from *necessitate de re*.

It seems to me that even if Adam does have such essential properties, there is no procedure at all for finding out what they are. And it also seems to me that there is no way of finding out whether he *does* have any essential properties. Is there really a good reason, then, for supposing that he does?

The distinction between essential and non-essential properties seems to be involved in one of the traditional ways of dealing with the problem of *knowing who*.⁸ If this way of dealing with that problem were satisfactory, then the doctrine of essential properties might have a kind of independent confirmation. But I am not sure that is satisfactory. The problem of *knowing who* may be illustrated in this way. I do not know who it was who robbed the bank this morning, but I do know, let us assume, that there is someone who robbed the bank and I also know that that person is the man who drove off from the bank at 9:20 A.M. in a Buick Sedan. For me to know *who* he is, therefore, it is not enough for me to have information enabling me to characterize him uniquely. What kind of information, then, *would* entitle me to say that I know who he is? The essentialistic answer would be: "You *know who* the bank robber is, provided that there is a certain set of properties E which are essential to the x such that x robbed the bank and you know that x has E and x robbed the bank." But if my doubts about essential properties are well-founded, this solution to the problem of knowing who would imply that the police, though they may finally "learn the thief's identity," will never know that they do. For to *know that one knows who* the thief is (according to the proposed solution) one must know what properties are essential to the thief; and if what I have said is correct, we have no way of finding out what they are. How are the police to decide that they know who the thief is if they have no answer to the metaphysical question "What are the essential properties of the man we have arrested?"⁹

⁸ Compare Aristotle, *De Sophisticis Elenchis*, 179 b 3; Petrus Hispanus, *Summulae Logicales*, ed. I. M. Bochenski (Turin, 1947), 7.41; Franz Brentano, *Kategorienlehre* (Leipzig, 1933), p. 165.

⁹ Hintikka says that we know who the thief is provided that there exists an x such that we know that the thief is identical with x (*op. cit.*, p. 153). But under what conditions may it be said that there exists an x such that we know that the thief is identical with x? Presumably, if ever, when we catch him in the act—when we *see* him steal the money. But the teller saw him steal the money and *she* doesn't know who he is. I have suggested elsewhere a slightly different way of looking at these questions; compare *op. cit.*, pp. 789–791, and "Believing and Intentionality," *Philosophy and Phenomenological Research*, XXV (1964), pp. 266–269, esp. p. 268.

It is assumed, in many writings on modal logic, that “Necessarily, for every x , x is identical with x ” implies “For every x , necessarily x is identical with x ,” and therefore also “For every x and y , if x is identical with y , then necessarily x is identical with y .” But is the assumption reasonable? It leads us to perplexing conclusions: for example, to the conclusion that *every* entity exists in *every* possible world and therefore, presumably, that everything is an *ens necessarium*.

Why assume that necessarily the evening star is identical with the evening star? We should remind ourselves that “The evening star is identical with the evening star” is not a logical truth, for it implies the contingent proposition “There is an evening star,” and that its negation is not “The evening star is diverse from the evening star.” Wouldn’t it be simpler to deny that “Necessarily, for every x , x is identical with x ” implies “For every x , necessarily x is identical with x ”? Then we could deny the principle *de dicto*, “Necessarily the evening star is identical with the evening star,” and also deny the principle, *de re*, “The evening star is necessarily identical with the evening star.”¹⁰ We could still do justice to the necessity that is here involved, it seems to me, provided we continued to affirm such principles, *de dicto*, as “Necessarily, for every x , x is identical with x ” and “Necessarily, for every x and y , if x is identical with y then y is identical with x ,” and such principles, *de re*, “The evening star, like everything else, is necessarily self-identical.”

¹⁰ I have discussed this possibility in “Query on Substitutivity,” in *Boston Studies in the Philosophy of Science*, Vol. II, ed., Robert S. Cohen and Marx W. Wartofsky (New York: The Humanities Press, 1965), pp. 275–278.

If we deny that “Necessarily, for every x , x is F ” implies “For every x , necessarily x is F ,” then presumably we should also deny that “It is possible that there exists an x such that x is F ” implies “There exists an x such that it is possible that x is F .” But isn’t this what we should do? One could hold quite consistently, it seems to me, that though it is possible that there exists something having the properties that Christians attribute to God, yet nothing that does exist is such that it is possible that *that* thing has the properties that Christians attribute to God.