THE PRINCIPLE OF THE IDENTITY OF INDISCERNIBLES: A FALSE PRINCIPLE

ALBERTO CORTES

Gottfried Wilhelm von Leibniz was the first person to state the Principle of the Identify of Indiscernibles (which we shall often abbreviate as PII) and he gave several formulations of it, some of them suggesting that the principle was a synthetic statement, and some suggesting that it was a logical principle. In any case, all the formulations given by Leibniz are stated in negative fashion: "that there are not in nature two indiscernible real absolute beings,"[1] or, "no two substances are completely similar, or differ *solo numero*."[2] In ordinary language, it is very difficult to state the principle in a positive manner. Roughly, it would be something like: "If two objects have all properties in common, then they are only one object." This formulation, or many similar to it, though stating the principle in a positive manner, has the disadvantage of being somewhat paradoxical, since it first entertains *two* objects only to deny that there are two.

It has been suggested by symbolic logicians that the principle should be stated, (using second-order quantificational theory with equality)

$$(x)\,(y)\,[(F)\,(F_x \leftrightarrow F_y) \rightarrow x = y] \qquad (1)$$

In words, for any object x and for any object y, if every property F of x is a property of y, and every property of y is a property of x, then x is numerically identical to y. Since 'x' and 'y' are variables which take as values elements of the domain, and both 'x' and 'y' can take the same value this formulation seem to present no problems as regards clarity.

As a matter of fact, the converse of (1), the Principle of *Indiscernibility of Identicals*,

$$(x)\,(y)\,[x = y \rightarrow (F_x \leftrightarrow F_y)] \qquad (2)$$

can be shown to be a theorem schema of first order quantificational logic with equality.[3] In (2), it is understood that F is an arbitrary predicate.

Many logicians believe that (1) is also a logical theorem (therefore a logical truth), and that therefore one can use a biconditional formed of (1) and (2) as a definition of the identity sign '='[4] (of *numerical identity* in terms of *properties*),

$$(x)\,(y)\,[x = y \leftrightarrow (F)\,(F_x \leftrightarrow F_y)]. \qquad (3)$$

Consider a proof of (1). Let me first present a very informal one and follow it by a modification of a formal proof shown to me by Professor Charles (Danny) Daniels.

Informal proof of (1); Assume that objects a and b satisfy the antecedent of (1),

$$(F)\,(F_a \leftrightarrow F_b). \qquad (4)$$

If now we instantiate 'x' to 'a,' 'y' to 'b,' and 'F' to the property of being numerically identical to a, by (4), b must also have that property. Therefore $b = a$.

For a formal proof of (1) I use a "natural" formal system like that of Kalish and Montague.[5] Carnap has a proof of this theorem too.[6]

Formal proof:

To prove:

$$(x)\,(y)\,[(F)\,(F_x \leftrightarrow F_y) \rightarrow x = y] \qquad (1)$$

prove,

$$[(F)\,(F_x \leftrightarrow F_y) \rightarrow x = y]$$

for arbitary objects x and y. Assume:

$$(F)\,(F_x \leftrightarrow F_y). \qquad (5)$$

Instantiate 'F' to the property $G = (\lambda z)\,(z = x)$, *being identical to* x. [Read, "the property (or class) of those z such that $z = x$]."[7] We obtain,

$$(\lambda z)\,(z = x)_x \leftrightarrow (\lambda z)\,(z = x)_y$$
$$x = x \leftrightarrow y = x$$

using theorem $x = x$ we have, $y = x$, therefore, $x = y$ for arbitrary x and y, and (1) is therefore a theorem.

Naturally, since I claim that PII is false, and the above proof seems to establish that PII not only is true, but it is a logical truth, I must show that the proof is fallacious. This I shall do in what follows.

Before we consider a critique of this proof I should show that in any given context both free variables and bound variables beyond the quantifier behave, with respect to their referential properties, just as individual constants do. This point aims at simplfying the presentation so that I shall not have to treat free *variables* and *names* separately; I shall not have to treat (4), (5) or any proof carried out within quantifiers separately. But limitations of time do not permit me to do this, and I shall assume it.

However, it might still be argued, as Quine has done, that in a logically austere language names are quite superfluous:[8]

> Chief among the omitted frills is the *name*. This . . . is a mere convenience and strictly redundant, for the following reason. Think of '*a*' as a name, and thing of '*Fa*' as any sentence containing it. But clearly '*Fa*' is equivalent to

'(Ex) (a = x · Fx).' We see from this consideration that 'a' needs never occur except in the context 'a=.' But we can as well render 'a=' always as a simple predicate 'A,' thus abandoning the name 'a.' 'Fa' gives way thus to '(Ex) (Ax · Fx),' where *the predicate 'A' is true solely of the object a*. [My italics.]

It may be objected that this paraphrase deprives us of an assurance of uniqueness that the name has afforded. It is understood that the name applies to only one object, whereas the predicate 'A' supposes no such condition. However, we lose nothing by this, since we can always *stipulate* by further sentences, when we wish, *that 'A' is true of one and only one thing.*

According to Quine, then, one can always eliminate a name by concocting an appropriate predicate which is true only of the object designated by that name.

But how is one to know that the predicate one has concocted does not apply simultaneously to more than one object? One cannot merely stipulate that it be so. One has to presuppose that if two objects are distinct they must not possess all properties in common. And this is nothing less than the contrapositive of PII. Therefore, it is evident that in order to eliminate proper names, *a la Quine*, one has to presuppose PII. Needless to say, we shall not admit such a presupposition in a critical analysis of PII. Therefore, from the point of view of this paper, proper names, or their logical equivalents (e.g., pronouns), are logically fundamental ingredients of both ordinary and formal languages.

Let us now take a very careful look at the proofs (informal or formal) of PII which we presented above. We assumed that there were objects a and b such that Equation (4), [or Equation (5) for arbitrary x and y]

$$(F)\,(F_a \leftrightarrow F_b) \tag{4}$$

was true of them; or, in other words, that a and b shared all properties on common. From that fact it followed that since a had the property of being numerically identical to itself, then, since b has the same property, it must also be identical to a. In symbols, $a = b$, which is what had to be proved.

Consider now some of the assumptions that we make within our conceptual or linguistic scheme whenever we make an identifying reference of a particular by means of a proper name (or of any functionally equivalent referring expression). First of all, as Strawson has correctly observed in his essay *Individuals*, it serves no purpose using a name for a particular unless the listener knows who or what is referred to by the use of the name, or unless one would be ready to describe who or what does the name designate. In spite of the obviousness of this remark, it is disregarded repeatedly by those who treat the problem of identity: consider the statement at the beginning of the proof above. One could hardly imagine a simpler and more general situation than the one in which all that has been assumed is

that there are objects a and b. We have not even assumed that there are two objects; there could be only one to which both names refer. Nevertheless, in such a general situation, if there are *two* objects, then 'a' refers to one of them and 'b: to the other! If I talk about a, I tacitly assume that a may be identifiable; and if I talk again about a, I tacitly assume that a may be reidentifiable. This means that conceptually I have individuated a and individuated b as soon as I assumed the seemingly contentless proposition that there were objects a and b. In other words, in a situation where there are things, *no matter how insignificant* my knowledge about them may be the mere fact that I use proper names 'a,' 'b,' etc., to talk about them implies that I regard them as being individuals, and hence discernible each from other things. And what does it mean to say that something is an individual and hence discernible? It means that such-and-such a thing is true of this object which differentiates it from any other object.

It appears, therefore, that in the seemingly neutral assumptions which precede the so called proofs of PII one is already presupposing that one is dealing with *individuals*. And it is analytically true that individuals obey PII since by definition no two individuals can share all properties.

Furthermore, it is not only in the assumptions which precede the proof of PII that individuals are assumed. They are assumed also whenever it is claimed of *something* that *it* has some self-referential property like self-identity since in such a claim there are presupposed conceptual identification and reidentification of such an object.

Another way of looking at these proofs of PII is to regard them as assuming that objects can be individuated by merely naming them! This is a *logical* mistake, since, as we have seen above, one can only refer to a particular if that particular is identifiable, and one can refer to it again only if it is reidentifiable; that is, if it is an individual.

Have I shown that PII is false? No, I have not. I have merely shown that proofs like the ones I presented, which use self-referential properties for individuating objects, are logically deficient. This step, of course, was crucial towards establishing my claim that PII is false, but it was not enough.

The following two claims, once established—and they can be established—seem to me to give a great deal of support to the claim that PII is actually false.

First, it seems that nature does provide us with entities which appear to violate PII. These are the fundamental particles called *bosons* by physicists, of which light quanta are a good example. In this case, the way to establish that bosons violate PII is inductive: one hypothesizes that bosons do violate PII and then observes whether the physical consequences of such a feature do give rise to observable phenomena. It is my claim that they do; that this is the way to interpret the quantum behavior of bosons.

Second, it has been claimed that there may be other properties that would individuate entities like bosons. One such property was suggested by Bas van Fraassen: he claims that the history of an object individuates that object.[9] It seems to me that the only way to deal with such properties is to treat them one by one and to show that they could not individuate otherwise identical particles. For the property of "having a unique historical past," I have shown elsewhere that it does not individuate otherwise qualitatively identical objects or particles.

NOTES

[1]Quoted by Bertrand Russell in *The Philosophy of Leibniz* (London: George Allen, 1951), p. 54.

[2]Ibid., p. 58.

[3]See for example Kalish and Montague, *Logic* (New York: Harcourt, Brace and World, 1964), pp. 222-23.

[4]Whitehead and Russell, *Principia Mathematica* (London: Cambridge University Press, 1910, 1964), Sec. 13; Rudolph Carnap, *Introduction to Symbolic Logic and Its Applications* (New York: Dover, 1958), Sec. 17; Alfred Tarski, Introduction to Logic (New York: Oxford University Press, 1941, 1966). Chapter III.

[5]Kalish and Montague, op. cit.

[6]Rudolph Carnap, op. cit., Sec. 17.

[7]Ibid.

[8]W. V. Quine, *Philosophy of Logic* (Englewood Cliffs, N.J.: Prentice Hall, 1970), Chapter 2.

[9]Bas van Fraassen, "Probabilities and the Problem of Individuation," in *Probabilities, Problems and Paradoxes*, ed. Sidney A. Luckenbach (Encino, Calif.: Dickenson Publishing Co., 1972).

Texas Tech University