

Impredicative Identity Criteria

LEON HORSTEN

University of Bristol

In this paper, a general perspective on criteria of identity of kinds of objects is developed. The question of the admissibility of impredicative or circular identity criteria is investigated in the light of the view that is articulated. It is argued that in and of itself impredicativity does not constitute sufficient grounds for rejecting a putative identity criterion. The view that is presented is applied to Davidson's criterion of identity for events and to the structuralist criterion of identity of places in a structure.

“...the first concern of the philosophy of any subject matter must be to enhance our powers of finding the elucidation [...] for its disputed identity questions.” [Wiggins, 1980, p. 53–54]

1. Plea for a General Perspective

In philosophical arguments, criteria of identity have continuously played a role at least since the work of Frege. Philosophers have sought criteria of identity for kinds of entities. They have argued for and against proposed criteria of identity, and have argued for and against the importance of criteria of identity. They have investigated how one ought in general to go about searching for a criterion of identity for a given kind of entities.

A general outlook on what acceptable criteria of identity are has to a large extent been lacking. One of the main objectives of the present paper is to develop such an outlook. I want to develop a general perspective on what identity criteria are. An account of on criteria of identity will be articulated which applies both to kinds of concrete entities and to kinds of abstract entities.

Such a theory will have to be somehow abstracted from concrete uses of criteria of identity in philosophical argumentation. The hope is that in successful philosophical uses of criteria of identity, a general

and robust concept of criterion of identity is implicit. It may turn out in the final analysis that there are no successful criteria of identity for kinds of concrete entities. Or it may turn out in the end that no illuminating *general* concept of criterion of identity can be abstracted from the manifold of successful uses of particular identity criteria. But only a philosophical investigation into what criteria of identity are might bear this out.

From a historical perspective, one can see why most philosophers in the twentieth century were not terribly interested in developing a systematic theory of criteria of identity in general. Frege already indicated in his crystal clear manner that there are certain things that criteria of identity do not deliver. Frege discussed the question of criteria of identity for natural numbers. As is well-known, he proposed what came to be known as *Hume's Principle*: the number of the *Fs* is the number of the *Gs* if and only if there is a one-to-one correspondence between the *Fs* and the *Gs*. He pointed out that his criterion of identity for numbers will leave open the question whether the number five, for instance, is Julius Caesar. In general, a criterion of identity for *K* will not decide whether a given *x* is a *K*. For this reason, Frege thought that criteria of identity should not be regarded as *basic* principles. An acceptable criterion of identity for *K* can only flow from a prior philosophical theory of *K*.

The logical positivists essentially agreed with Frege. They wanted philosophical theories to explicate how objects and properties are constructed from our sensory experience. So for them, too, criteria of identity could serve at best as intermediate steps in the construction of satisfactory philosophical theories.

Since the demise of logical empiricism it is no longer unconditionally demanded of philosophical theories that they explicate how their objects of investigation are constructed. Quine maintained that philosophical theories must nevertheless pass a minimal test. At a minimum, they must separate the elements of their domain of investigation. In short: they must entail a criterion of identity for their subject matter. This is the content of his famous slogan: *no entity without identity*. This requirement was most forcefully used as a critical tool. In the 1950s and 1960s the semantics of modal logic was formulated in terms of *intensions*. Quine conjectured that no satisfactory criterion of identity for meanings of sentences (intensions) could ever be found. And if this is so, then we should, according to Quine, have no reason to think that a subject matter of this theory exists. In this way, Quine licensed the inference from the absence of criteria of identity to scepticism about the existence of the entities that are investigated.

Since then, philosophers have argued that it is unreasonable to expect that for *all* kinds of entities, criteria of identity are to be found [Strawson, 1976]. Today, few philosophers are even willing to insist with Quine on criteria of identity for all *scientific* subject matters. A satisfying criterion of identity for algorithms, for instance, is notoriously hard to come up with. Yet few will maintain that *therefore* our theories of the notion of algorithm fail to meet minimal philosophical standards.

These contemporary views seem to me justified. But they provide no good reasons for thinking that a general view of criteria of identity cannot be had or is not worth having. Currently, the requirements for philosophical theories are taken to be no different from the requirements on scientific theories in general. A philosophical theory must articulate principles which, when combined, yield philosophical insight into the subject matter under investigation—that is all. But criteria of identity can surely contribute greatly to this: successful criteria of identity considerably enhance our insight in the nature of kinds of objects.

So having a general perspective on what criteria of identity are is important not only in its own right, but also for methodological reasons in philosophy. It is important to be able to reliably judge when we are in possession of a successful criterion of identity for a kind of objects. Again, Quine famously held it to be a necessary condition of adequacy for philosophical theories that they entail criteria of identity for the kinds of objects that constitute their subject matter. But in the absence of a general perspective on criteria of identity, it is difficult to assess in specific instances whether this demand is justified. Williamson endeavored to uncover general methods for finding a criterion of identity for a given kind of objects [Williamson 1986], [Williamson 1990]. Lowe and Williamson debated about the form that criteria of identity in general take [Lowe 1991], [Williamson 1991]. And there was a discussion between Quine and Davidson on the admissibility of some form of circularity in criteria of identity. In all these issues and disputes, one would expect the positions to be based on a general view of what criteria of identity are.

But such views are rarely made explicit. The only explicitly articulated view of which I am aware, is that of Lowe [Lowe 1989b]. While I find much merit in Lowe's view, the perspective developed in this article differs considerably from his. The divergence of opinion is located chiefly in the treatment of impredicativity of identity criteria. It will emerge that this has implications for the assessment of certain particular criteria of identity that have been proposed in the literature.

2. The Form of a Criterion of Identity

It is held that a criterion of identity is an answer to the question “when is a thing x (of sort K) identical to a thing y (of sort K)?” If that is so, then it seems that the logical form of a criterion of identity is:

$$\forall x \forall y \in K : x = y \leftrightarrow \Phi(x, y)$$

The paradigmatic, oft-cited example here is the principle of extensionality in set theory.

It has been argued that not all requests for a criterion of identity take the above-mentioned form. For such a request can also take the form: “when is the f of a thing x equal to the f of a thing y ?” Here it is assumed that there is an underlying domain of entities D available, and a (definable) function f such that $f(D)$ constitutes a kind. Criteria of identity that are supplied in response to such questions take the form:

$$\forall x \forall y \in D : f(x) = f(y) \leftrightarrow \Phi(x, y)$$

The paradigmatic example here is Frege’s criterion of identity for directions of lines in terms of parallelism between lines. Criteria of the first kind are called *level 1 criteria*; criteria of the second kind are called *level 2 criteria*.

Some philosophers have wondered which of these two is the more fundamental type. And some have wondered whether one of these types can somehow be reduced to the other [Williamson 1990, p. 146–147], [Lowe 1989b, p. 6], [Lowe 1991], [Williamson 1991]. Without going into details here, I take the outcome of this debate to be the following. Judging from examples of successful or at least promising criteria of identity, there is *prima facie* evidence that criteria of identity can take these two forms. And while formal reductions from one format to the other are possible, they seem artificial. At any rate, we should not be too preoccupied with logical form, for there may be criteria of identity which do not clearly fit into either of the types discussed so far. The form of criteria of identity over time, for instance, may be the following [Merricks 1998, p. 116]:

$$\forall x, y \in S \forall t \forall t' : x \text{ at } t = y \text{ at } t' \leftrightarrow \Phi(x, y, t, t')$$

If criteria of identity have this form, then they are not clearly of level 1, nor are they clearly of level 2.

There appear to be clear examples of kinds which *invite* level 1 criteria: events, classes, material objects... And there appear to be clear examples of kinds that invite level 2 criteria: meanings (of sentences),

bird species (to which concrete animals belong), chemical substances (to which laboratory samples belong),... So it shall be assumed that criteria of identity come in at least two flavors (level 1 and level 2), perhaps in more. In what follows, I shall be mainly concerned with level 1 criteria. But matters of form of criteria will not be regarded as terribly important.

3. Plea for a Theoretical Conception of Identity Criteria

The literature shows no consensus on what *kind of principle* a criterion of identity is. Broadly speaking, three answers have been given:

1. Criteria of identity are semantical principles.
2. Criteria of identity are epistemic principles.
3. Criteria of identity are metaphysical principles.

Intermediate positions are possible. For instance, it has been claimed that criteria of identity are semantical-cum-metaphysical principles. This discussion is not inconsequential. It will be argued that it is intertwined with discussions about the adequacy of particular criteria of identity.

According to the view that identity criteria are primarily semantical principles, criteria of identity are about sameness and difference of reference of simple and complex names. That criteria of identity are *primarily* semantical principles is argued in [De Clercq 2005, p. 23, 32]. However, he does not deny that criteria of identity also have a metaphysical component [De Clercq 2005, p. 25]: at one place he labels them as semantic-metaphysical principles. In his earlier work, Lowe likewise describes criteria of identity as metaphysical-cum-semantic principles [Lowe 1988, p. 62–63], or even as plainly semantical principles [Lowe 1989b, p. 13]. Strawson can be regarded as a defender of the view that identity criteria have a basic epistemic component. This follows from his repeated insistence that identity criteria should be *applicable* [Strawson, 1976, p. 39, 49, 50], [Strawson, 1997, p. 2]. Williamson argues that criteria of identity are first and foremost metaphysical principles [Williamson 1990, chapter 9]. In his more recent work, Lowe also expresses this view [Lowe 1998, p. 36, 44]. The position that is developed in this paper will be closer to that of Lowe (in his later work) and Williamson than to that of Strawson.¹

¹ I do not claim that Williamson or Lowe would accept my version of the metaphysical interpretation of identity criteria.

There is *prima facie* evidence that criteria of identity are metaphysical principles. An identity principle for a sort *K* expresses under which conditions an *x* and an *y* of the sort *K* are identical. The conditions are likely to contain only metaphysical notions such as ‘sort’, ‘place’, ‘part’, ‘stage’, Hence the criterion must, if it is true, hold with metaphysical necessity. It seems, therefore, that a criterion of identity for *K* must then flow from metaphysical facts about the nature of *K*.

If this is so then it seems that one can agree, with Lowe, that [Lowe 1998, p. 36]:

[...] identity criteria are precisely *metaphysical* principles, telling us (as Locke would put it) what identity *consists in* for objects of given kinds.

But such talk has to be used and interpreted with extreme care. For it is sometimes read as implying that an identity criterion is supposed to tell us what *makes* an object *x* equal to (or different from) an object *y*. One finds such an interpretation of Lowe’s stance in [Noonan 1988, p. 80–81]. Everyone agrees—or should agree—that such an *analysis* of the relation of identity of objects of a kind can never be had. Jubien even goes so far as to say that identity criteria for kinds of objects can *only* coherently be interpreted as giving an analysis of the relation of identity [Jubien 1996, p. 347]. But an analysis of a relational concept must explicate the meaning of the concept in terms of more primitive concepts. If this is what is required, then there are no criteria of identity at all: even the principle of extensionality is *not* an acceptable criterion of identity for classes [Jubien 1996, p. 348]. Surely this interpretation of what criteria of identity are is not forced upon us. Given the fact that identity is a primitive relation, such a theory cannot give an analysis of identity in terms of more basic concepts. But it need not do this, just as a philosophical account of a property or relation need not restrict itself to an analysis of the natural language meaning of an expression that refers to the property or relation [Anderson 1993]. A successful identity criterion for a kind should provide no more and no less than a good *theory* of identity for objects of the kind in question. It should link the relation of identity for *Ks* to other concepts without abandoning the thesis of the primitiveness of identity.

Moreover, it should be stressed that at the outset there is no reason to believe that a criterion of identity tells us *all* there is of metaphysical interest to know about the conditions under which an object *x* and an object *y* of a kind *K* are the same. For all the metaphysical view is committed to, there may be two coextensive but intensionally distinct

criteria of identity for some one kind of objects, each explicating *aspects* of what “identity consists in” for the kind under investigation.

If criteria of identity were semantic principles, then one would expect them to typically contain semantic notions, such as ‘meaning’, ‘reference’, ‘truth’, ‘synonymy’, But they typically don’t. In particular, the left-hand-side of a criterion of identity does not involve presentations of objects. It specifies the conditions under which an *object* x (of sort K) is identical with an *object* y .

It is sometimes thought that if you put it this way, then it is hard to see how an identity criterion could ever be *informative* [De Clercq 2005, p. 23–24]. For it seems that we already know all there is to know about identity of objects: every object is identical with itself and different from every other object.

But this problem is easily solved. The semantical value of a criterion of identity emerges in *instantiations* that are made of it. As soon as the initial universal quantifiers of a criterion of identity are instantiated, presentations of objects (simple or complex names) come to the fore. And then the left-hand-side of the equivalence becomes a nontrivial *proposition* which is true if and only if the closed terms in question refer to the same object. In this way, the instantiated criterion implies information about co-referentiality of closed terms. So it is true that criteria of identity have fairly immediate semantical implications: they can typically be used to determine sameness of reference of individual expressions. But that alone does not make them semantical principles.

As with the semantical interpretation of identity criteria, one could argue that if criteria of identity were epistemological principles, then one might expect them to contain epistemic notions such as knowledge, justification, But according to Strawson this is the wrong way to picture the epistemic nature of identity criteria. The epistemic nature of identity criteria is betrayed by their *point*, by their function. An identity criterion is something to be *applied*. In principle, an identity criterion for K must in his view give us a decision method for identity questions for members of K . Naturally, Strawson will concede that identity criteria must give us decision procedures only in a highly idealized sense. For the right-hand-side of an identity criterion may (and typically will) contain highly non-observational vocabulary, and apart from that there might be computation-theoretic reasons why a decision procedure in the Church-Turing sense does not exist for the criterion. Indeed, it will contain metaphysical vocabulary and to this extent identity criteria are *also* metaphysical principles.

There surely is something in Strawson’s claim. It seems true that the core meaning of the word ‘criterion’ contains, beside a metaphysical component, such an epistemic component: a criterion just *is* something

that is meant to be applied. But the metaphysical component seems just as real. Williamson puts it thus [Williamson 1990, p. 148]:

The word ‘criterion’ in philosophy has justly fallen under a cloud. For it was used with the effect of conflating two quite different questions, the metaphysical ‘What is it for this to be so-and-so?’, and the epistemic ‘How can we know that this is so-and-so?’. The equivocation has been observed still active in the phrase ‘criterion of identity’.

Conceptual purity in philosophy is a nice thing. If we want the concept of identity criterion to do clean work, the epistemic and the metaphysical component should not be run together in one notion. Therefore I advocate that we follow Williamson in considering the metaphysical interpretation of the notion of identity criterion to be the more basic one [Williamson 1991, p. 149], and take the requirement of applicability to be one that some identity criteria do not satisfy and others do. We shall call this latter kind *applicable* or *predicative* identity criteria.

Admittedly this is a plea for (slight) conceptual reform but in this case it is beneficial. Let those who want to reserve the name ‘criterion of identity’ for the curious hermaphroditic thing that it really is in ordinary language coin a new name for the metaphysical concept that I have tried to elucidate. Perhaps the name ‘*principle* of identity’ would fill this bill. However, since there is every reason to believe that this proposal for terminological reform will simply be disregarded by the philosophical community, I shall continue to use the term ‘criterion of identity’.

In sum, if we must choose between the three interpretations listed at the beginning of this section, we should see identity criteria as metaphysical principles. But perhaps a broader interpretation of the notion of identity criterion is in order. If metaphysics is the science with the widest possible subject matter, then it seems that certain identity criteria are most appropriately allocated to the special sciences. Consider again the intuitive level 2 identity criterion for bird species:

The species to which bird x belongs is identical with the species to which bird y belongs if and only if x and y are able to interbreed or (in case x and y are of the same sex) there exists a bird z which can interbreed with x and with y .²

² As it stands, this identity criterion is of course defective. For one thing, the interbreeding relation is not transitive. For this reason, it is argued in [De Clercq & Horsten 2005] that an acceptable identity criterion for bird species will be an equivalence-approximation of the intuitive criterion that is formulated here.

It seems more proper to classify this as a *biological* identity criterion than as a metaphysical principle. And in a similar vein identity criteria for meanings of sentences *can* properly be said to be semantical principles. In general, it seems most fruitful to interpret criteria of identity as *theoretical principles*. In order to be acceptable, a criterion of identity should then hold with theoretical necessity (metaphysical necessity, biological necessity,...), for they should flow from theoretical (metaphysical, biological,...) principles. In the sequel, I shall be mostly concerned with identity criteria which do not belong to any of the special disciplines, i.e., I shall be mostly concerned with metaphysical criteria of identity.

4. The Metaphysical Role of Criteria of Identity

If the suggestion made here is seconded and we distinguish between applicable (predicative) and non-applicable (impredicative) criteria of identity, then it of course becomes an important question what *makes* a criterion of identity applicable. I will set this question aside for the moment, and return to it later. For on the proposed conception of identity criteria, there is a prior question: what is the *role* of criteria of identity in metaphysics? This is not the same as asking what the epistemological or semantical *use* of identity criteria consists in. For these questions have already been partially answered: criteria of identity determine and in many cases allow us to decide questions about identity and difference of denotation of closed expressions.

There is a consensus that a criterion of identity must be *informative*. An identity criterion must not be a truth of logic.

Let us—without loss of generality—disregard the meaning postulates for the moment and—hopefully still without loss of generality—focus on level 1 criteria. Such a criterion is informative only if it excludes certain interpretations of the non-logical vocabulary occurring on the right-hand-side. Since moreover criteria of identity must derive from (or indirectly express) aspects of the nature of the kind under investigation, it follows that successful identity criteria provide structural ontological insight into the kind of entities in question. Often this is effected by a *sharpening of the concept* that one is trying to capture—I am using ‘kind’ and ‘concept’ interchangeably here. The axiom of extensionality makes it clear that an extensional concept of class is aimed at. Hume’s principle makes it clear that it is not the ordinal concept of number that is intended.

More often than not, criteria of identity are proposed and then on closer scrutiny prove to be deficient. Even repeated such failures need not—*pace* Quine—in itself be cause for scepticism about the kind of entities in question. Often, the reverse is rather the case: deeper insight into the kind of entities under investigation is acquired.

As a first example, consider the decades-long hunt for a criterion of identity for propositions or meanings of sentences. Despite heroic efforts, we today still lack a satisfactory criterion of identity for propositions. But many a philosopher of language takes this search to have revealed a deep fact about propositions, viz. that there is an almost one-to-one structure-preserving correspondence between sentences and propositions: “propositions are fine-grained.”

As a second, more elaborated example, consider the problem of finding a viable criterion of identity for material objects. There is a straightforward candidate: coincidence in space at a moment in time. Counterexamples were near at hand: the statue and the lump of clay. In fact, the example of the statue and the lump of clay appeared to show that it is possible for two objects to coincide in space at *all* moments in time. Strawson then proposed (and dismissed) the following identity criterion for material objects [Strawson, 1976, p. 39–40]:

$x = y$ if and only if x and y belong to the same kind of material object and at some moment in time x and y spatially coincide.

Actually, Strawson required for x and y belonging to the same kind of material object to be identical that they spatially coincide at *all* times. But, as De Clercq notes, this is equivalent to them spatially coinciding at at least one moment in time. This criterion goes back to [Locke 1689, chapter 27, section 3], so let us call it *Locke’s criterion*.

This proposal was countered by the letter example of Kit Fine. Consider a letter written by a man to his lover, in which he announces his intention to break off their relationship. She returns the letter to him after writing on the back: “Here you have your horrible letter back.” So we have two letters that coincide in space at some times. (It is not hard to see how certain superficial objections can be dismissed by finessing Fine’s example.)

De Clercq defends Locke’s criterion in [De Clercq 2005]. He reacts to this example by saying that either the two letters do not belong to the same (substantial) kind, or, contrary to first appearances perhaps, there is only one letter. But an alternative reaction to the letter example would be to deny that letters—and meaning-carrying entities in general—are material objects. If this is the correct response, then we have gained new insight in what it is to be a material object. Alternatively, this latter reaction can be seen as a sharpening of the concept of material object. We do not need to feel forced to take letters to be abstract objects. A suggestion would be that aside from material and abstract objects, we have a third kind: symbolic objects.

In sum, criteria of identity separate the objects of the domain of discourse by means of theoretical relations. Thereby both successful and many demonstrably unsuccessful identity criteria furnish ontological insight into structural possibilities and impossibilities concerning kinds of entities.

Now one can wonder *why* criteria of identity are metaphysically useful in this way. Why do they typically function as imposing constraints on metaphysical theories of kinds of objects? The reason is, I think, that the possibilities and impossibilities that are claimed by identity criteria lend themselves to being critically evaluated by our metaphysical intuitions or common sense. They lie closer to our intuitions than the basic principles of high level metaphysical theories with which they are connected.

But Frege's point remains. Criteria of identity can never be a substitute for a good philosophical theory of the subject matter under investigation. And if we regard them as theoretical principles, then criteria of identity need not even give us an idealized decision procedure for identity questions concerning objects of the kind in question.

5. Adequacy Conditions

Based on the general view of criteria of identity that was articulated in the previous sections, adequacy conditions for criteria of identity can now be listed. This will enable us to judge in concrete instances whether a proposed criterion of identity for a kind of entities is acceptable. The following is intended to be a list of individually necessary and jointly sufficient conditions for being an acceptable identity criterion:³

1. A criterion of identity must be *formally adequate*. By this is meant that the right-hand-side of a criterion of identity must be an equivalence relation.
2. A criterion of identity must be *materially adequate*. This is a correctness condition. It means that the criterion of identity must be a *true* principle.
3. Criteria of identity must be *necessary* truths. In order to be significant, they must follow from theoretical principles concerning the subject matter in question.

³ Note added in proof. In the meantime Hannes Leitgeb has shown that the above conditions are not jointly sufficient for acceptability [Leitgeb 2010]. In order to be acceptable, an identity criterion must in addition be natural in a way that is admittedly difficult to make precise.

4. A criterion of identity must be *informative*. An identity criterion should not be a truth of logic.

Whether a putative identity criterion is formally adequate is in most instances uncontroversial: it can usually be checked in a straightforward manner. The question whether a criterion is materially adequate is frequently debated on the basis of putative counterexamples, as is illustrated by the discussion of Locke's criterion in the previous section.

The reason that identity criteria must be theoretically necessary is, as we have seen, that their truth must flow from theoretical insight into the nature of the kind under consideration. For such theoretical insight consists of theoretically necessary truths. In philosophy, it is mostly the notion of metaphysical necessity that is relevant here.

The informativeness requirement is an expression of the fact that a criterion of identity must impose ontological constraints on the nonlogical relations and properties in terms of which it is formulated. This is the reason why Leibniz's principle according to which objects are judged to be identical if they have the same properties is generally taken to fail as a criterion of identity: it is a law of second-order logic. Thus informativeness is a matter of excluding models. This in turn means that informativeness comes in degrees. If one identity criterion excludes more models than another identity criterion, then the former is more informative than the latter.

It may be objected at this point that some hold that certain perfectly acceptable identity criteria are logical truths. In particular, neo-logicist views regarding Hume's principle and of the extensionality axiom for classes come to mind here [Hale & Wright 2001]. It would take us too far afield to enter into the debate about the viability of neo-logicist views with respect to the theory of the natural numbers and parts of set theory. But it may be observed that neo-logicism has hitherto remained a minority view. Indeed, Wright himself has in recent years sought to qualify his earlier claim that Hume's principle is a truth of logic.

The reader will have noticed that Strawson's requirement of *applicability* or predicativity does not occur in the list of requirements. The epistemic notion of applicability is desirable, to be sure, but according to the metaphysical conception of identity criteria it is to be treated as a supererogatory virtue.

6. Circularity

It is often retorted that identity criteria must fulfill this extra demand: they must not be circular or impredicative as it is sometimes called

[Lowe 1998, p. 45], [Quine 1985, p. 166], [De Clercq 2005, p. 26]. In this section, we discuss this requirement.

For level 2 criteria, there is in practice usually no danger of circularity. This is because the range of the function f in the criterion is mostly kept disjoint from the domain of the function.⁴

For level 1 criteria, the situation is somewhat different. In the literature, Davidson's criterion for identity of events is regarded as a paradigmatic example of a circular identity criterion. In this respect, it is contrasted with the extensionality criterion for classes, which is regarded as non-circular. We shall now discuss Davidson's criterion in some detail. But we shall also look at another identity criterion, which has received some attention in the literature on the philosophy of mathematics recently, and which suffers from the same sort of circularity. It is a structuralist criterion of identity for places in a structure.

Davidson's criterion says that an event x is identical to an event y if and only if x and y have the same causes and effects [Davidson 1970, p. 306]. Davidson has to my knowledge never said much about the *motivation* for his identity criterion for events. Perhaps its motivation was a metaphysical view, namely the view that events are fundamentally relational by nature. This could be a motivation for a principle of ontological economy (Ockham's razor), which takes the form of an identity criterion: if an event x and an event y play the same role in the cosmic causal network, then they are one and the same event. Even though this seems the best motivation that can be given for the criterion, it appears out of character with Davidson's general philosophical outlook, which suggests that he regarded events as concrete particulars in a more straightforwardly realist manner. So perhaps Davidson simply thought that the correct identity criterion for events is likely to be analogous to the principle of identity for classes (extensionality), except that there are also the effects that have to be taken into account.

However one seeks to motivate Davidson's criterion, it is generally regarded as circular, and this circularity is regarded as problematic. The following is a typical judgement about the criterion [Cleland 1991, p. 230]:

[T]he circularity involved in Davidson's proposal has to do with the fact that it individuates events only if they are already individuated. As a result it fails to provide adequate identity conditions for events.

⁴ The range and domain of the function f are not *always* kept disjoint: think of Hume's Principle.

Lowe brings out the circularity somewhat more precisely as follows [Lowe 1989a, p. 180–181]:

The reason why [the extensionality principle] escapes the [circularity] charge is that, *supplemented by the other axioms [of Zermelo-Fraenkel set theory]*, [the extensionality principle] serves to guarantee that there is just one empty set at the base of the hierarchy of sets and that any identity question concerning sets can ultimately be settled by reference to the empty set through related applications of [the extensionality principle]. [...] The real trouble with [Davidson's criterion], then, is that it is presented to us without the supporting framework of an axiomatic set theory of events comparable to that provided for [the extensionality principle] by axiomatic set theory. Nor does it seem at all likely that such a theory of events could realistically be developed, for no event seems apt to play a role analogous to that of the empty set [...]

One can see what Lowe probably has in mind here. In order to decide whether a set a is identical with a set b , the extensionality principle requires us to compare their respective elements. But this enquiry can only yield unambiguous results if it is known, for each $c \in a$ and $d \in b$, whether $c = d$. In order to decide these questions, we must compare the elements of c with those of d , and so on. The Foundation Axiom of set theory is equivalent to the statement that there are no infinite descending \in -chains. So our decision procedure must terminate. The terminating nodes of this procedure are of course comparisons of the empty set with other sets. If one works not in ordinary set theory but in Zermelo-Fraenkel set theory *with Urelements*, then the difference questions do not bottom out in the empty set.⁵ Instead, the decision procedure will lead to questions whether $u = v$ where u and v are Urelements. Thus, in set theory with Urelements, extensionality is predicative only if we have in addition a predicative identity criterion for the class of Urelements.

But Lowe's observations raise a question. If there were a first cause (a *Big Bang* event), and if the causal structure of the universe somewhat resembled the universe of sets, then would not Davidson's criterion be just as predicative as the principle of extensionality? Indeed, do not the above considerations suggest that whether Davidson's criterion is circular depends on the causal event structure of the world?

Quine and Strawson are more categorical than Lowe in their rejection of Davidson's criterion [Strawson, 1976, p. 39, 49–50], [Quine 1985, p. 166].

⁵ This was pointed out by an anonymous referee.

Strawson argues that Locke's criterion that was discussed earlier is unacceptable for the same reason [Strawson, 1976, p. 39]:⁶

But there remains something profoundly unsatisfactory about [Locke's criterion]. It seems that in order to apply it we must already be operating a principle of identity: for how else could we be sure that we had the identical individual, *a*, in all those positions in which we are then to ask whether we had, at the same times, the individual, *b*?

In his argument against Davidson's criterion on account of its circularity, Quine compares circularity of criteria of identity with impredicativity of *definitions* [Quine 1985, p. 166]:

For my own part, I welcome impredicative definitions. I have remarked that there is nothing wrong with identifying the most typical Yale man by averaging measurements and tests of all Yale men including him. But we now observe that impredicative definition is no good in individuation. Here a difference between the impredicative and the predicative emerges which is significant quite apart from any constructivist proclivities. We can define impredicatively but we cannot individuate impredicatively.

Davidson accepted Quine's objections and abandoned his criterion [Davidson 1985, p. 175]. But only when criteria of identity are regarded as by their very nature first and foremost *epistemic* principles do Quine's considerations have any force.

Let us pursue Quine's comparison with impredicative definitions a little further. In a famous passage, Gödel defends use of impredicative definitions as follows [Gödel 1944, p. 127–128]:

[...] it seems that the vicious circle principle [...] only applies if the entities involved are constructed by ourselves. In this case there must clearly exist a definition (namely the description of the construction) which does not refer to a totality of things to which the thing to be constructed itself belongs. If, however, it is a question of objects that exist independently of our constructions, there is nothing in the least absurd in the existence of totalities which can be described (i.e., uniquely characterized) only by reference to this totality.

Gödel's view of impredicative definitions today is the received point of view. If his view is correct and criteria of identity are conceived as metaphysical principles (as was urged in section 3), then it would seem that Gödel's remarks should also apply to identity criteria. There

⁶ Lowe's position on Locke's criterion is substantially the same as that of Strawson [Lowe 1988, p. 67].

would appear to be nothing in the least absurd in the existence of kinds of objects for which only circular identity criteria exist. If a criterion would somehow create the objects of the sort under investigation, then circular criteria would be problematic. But since the objects exist independently of the criterion, some criteria may well *have to be* circular. There is a circular structuralist criterion of identity for natural numbers: natural numbers are identical if and only if they occupy the same position in the natural number structure. It seems a perfectly good criterion. There is also a criterion of identity for natural numbers which is not circular: Hume's Principle.⁷ But *that alone* does not make the structuralist criterion deficient. For some kinds of entities—events, perhaps?—it may be that the only satisfactory criteria that exist are in some sense circular.

It might be retorted in defense of Quine that definitions are semantical principles, whereas on the view that is defended here identity criteria are metaphysical (or theoretical) principles. What goes for semantical principles does not necessarily go for metaphysical principles. But definitions must be backed up by existence assumptions. Before introducing the description term 'the most typical Yale man', it is incumbent upon us to establish that the class of Yale men is not empty; we can only impredicatively define the closure of an inductive operation on classes of numbers if the full (impredicative) second-order comprehension axiom is at our disposal. It is precisely these existence assumptions that opponents of impredicative definitions object to. So in the final analysis, the discussion about the admissibility of impredicative definitions is a metaphysical debate after all.

Ultimately, it would seem that the burden of proof lies with Quine and his followers. If predicativity is listed among the adequacy criteria for identity criteria, then good reasons have to be supplied for doing so. Until now, such reasons have not been forthcoming.

It is evident that circularity in Lowe's sense does not necessarily prevent a criterion from being informative. Davidson's criterion may be circular. But it imposes restrictions on the structure of the underlying event structure. So it is clearly informative. If the theoretical interpretation of identity criteria is adopted that was explicated in section 3, then impredicativity is a matter of *applicability* (in Strawson's sense) of the criterion. And applicability is an *extra* (epistemic) virtue that some

⁷ The way in which identity questions for numbers are grounded in the context of Hume's Principle is analysed in [Leitgeb 2009]. See also [Horsten and Leitgeb 2009]. Even though Hume's Principle is predicative as a criterion of identity in the sense explained earlier in this paper, its *deployment* by Frege was impredicative in another sense. The derivation of Frege's theorem from Hume's Principle relies on the impredicative second-order comprehension scheme [Linnebo 2004].

acceptable identity criteria have and others lack. Davidson's criterion may lack this virtue.

So far, we have been assuming that Quine, Strawson and Lowe are right in claiming that Davidson's criterion is circular. But *is* it circular? I shall argue that it depends on what our universe is like.

The way in which Davidson's criterion is circular is less straightforward than is appreciated in the literature. In order to bring out the subtlety of the impredicativity of Davidson's criterion, we have to describe the intended sense of impredicativity more precisely than is commonly done. Nevertheless, it is intended that the account which I am about to articulate is faithful to or at least compatible with the conception of impredicativity of identity conditions that is described in [Lowe 1989a].

A full-fledged mathematical analysis of the impredicativity of Davidson's criterion of identity criteria is left for another occasion. The present discussion is restricted to giving a precise description of the notion of impredicativity that I take to be operative, and to illustrating it on the basis of some examples.

If we want to be precise, then we must first introduce some terminology.

A *causal event structure* is a set of events e_1, e_2, \dots on which a causality relation C is defined. So, formally, an event structure is a *directed graph* with loops. In the sequel, we shall restrict ourselves to *finite* directed graphs.

It is a disputed question whether the causality relation must always be transitive or not—Davidson does not say. Here we do not impose a requirement of transitivity. We shall also impose no a priori anti-symmetry condition on C . Indeed, we do not want to exclude by fiat that two events causally influence each other, as it is the case in Newtonian action at a distance. To conclude, we shall not exclude that some events at least in part cause themselves, i.e., that are at least in part *causa sui*.

An event structure is Davidson-admissible if it satisfies Davidson's criterion; otherwise it is Davidson-inadmissible. In the sequel, Davidson-admissible structures will often be simply referred to as *Davidsonian structures*. The event structure $\{e_1, e_2\}$ with the causality relation being the empty relation is evidently not Davidsonian. But the event structure with the same underlying set of events but with the causality relation described as $e_1 C e_2$ is a Davidsonian structure.

As it will turn out, some Davidson structures are *applicable* or *predicative* or *non-circular* in the Strawson-Lowe sense, others are not. Our immediate task is to explain what it means for an event structure to be applicable.

As we have seen, the impredicativity of Davidson's criterion is an epistemological notion. Suppose we have a cognitive agent, who is given a causal event graph. For the agent, the vertices in the graph are *event descriptions* [Davidson 1970, p. 295].⁸ The agent holds open the possibility that there different vertices (i.e., pairs of event descriptions) that represent one and the same event.

And let us say that the agent knows that the causal event graph that is given to her is *sound*: if there is an edge from one event description e_1 to another event description e_2 , then the event denoted by e_1 is indeed a cause of the event denoted by e_2 (and the latter is an effect of the former).

Now suppose that we also assume that the graph that is given to the knowing agent is also *complete*. This means that each event is named by at least one event description and the denotatum of an event description e_1 causes the denotatum of e_2 if and only if the event description graph contains an edge from e_1 to e_2 . Then the knowing agent can in principle decide, on the basis of Davidson's identity criterion, for every pair of event descriptions e_1 and e_2 whether they denote the same event. Indeed, suppose that the knowing agent has a sound and complete causal finite event graph and Davidson's criterion of identity for events holds. Then the agent can algorithmically decide for every pair of event descriptions e_1 and e_2 whether they are co-designating. That this is so, is immediate: two event descriptions are co-designating if and only if the vertices representing them in the graph are causally connected in the same way to the rest of the graph. So if the agent has sound and complete information about the causal structure of the world, then *all* identity questions can be settled, using Davidson's criterion, on the basis of causality information alone. No information about the identity of denotations of event descriptions is needed. So in this sense, Davidson's identity criterion is *not* circular.

A similar conclusion can be reached for the principle of extensionality. Suppose that the agent has sound and complete information about the \in -relation. And suppose that there is a clear violation of the Axiom of Foundation. To keep matters simple, suppose the agent has two names of sets a and b , and the full information about the \in -relation is given to him by:

$$a \in b$$

$$b \in a$$

⁸ See also [Anscombe 1979].

$$a \in a$$

$$b \in b$$

Then the extensionality principle allows her to infer that $a=b$. In sum, if the agent has sound and complete information, then extensionality is not circular, even in the presence of an Anti-Foundation Axiom.

So this is presumably not the notion of circularity that those have in mind who accuse Davidson's criterion of circularity. To capture the notion of circularity that the objectors have (dimly) in mind, we must suppose that the causal event graph is not complete but *semi-complete*, in the following sense. As before, each event is named by at least one event description. In addition, for every event description e_1 and every event e , if the event denoted by e_1 is a cause of e , then there is an edge in the graph from e_1 to at least one event description denoting e . Intuitively, this means that the agent has "almost complete" causality information: for each event description e_1 , all the causal relations of its *denotatum* are represented in the structure description. But the causal relations are not necessarily represented under all presentations of the events.

I submit that this is the type of situation which those who accuse Davidson's criterion of circularity have (dimly) in mind. We shall see that many sound and semi-complete event description graphs are epistemically circular or *impredicative*.

The impredicativity of a davidsonian graph can be brought out as follows. Suppose e_1, e_2, e_3, e_4 are different vertices of a davidsonian graph $G = \langle V, C \rangle$. What we want to describe formally is the way in which the difference between the events represented by two vertices e_1 and e_2 *supervenes* on the difference between the events represented by two other vertices e_3 and e_4 , according to Davidson's identity criterion.

We shall say that the ordered pair $\langle e_1, e_2 \rangle$ immediately supervenes on $\langle e_3, e_4 \rangle$ if and only if $e_3 C e_1$ and $e_4 C e_2$, or $e_1 C e_3$ and $e_2 C e_4$. But the supervenience relation is intended to be order-independent. So we need a definition of immediate supervenience for unordered pairs. We say that the *set* $\{e_1, e_2\}$ immediately supervenes on $\{e_3, e_4\}$ if and only if:

- $\langle e_1, e_2 \rangle$ immediately supervenes on $\langle e_3, e_4 \rangle$, or
- $\langle e_1, e_2 \rangle$ immediately supervenes on $\langle e_4, e_3 \rangle$.

The notion of (possibly non-immediate) *supervenience* is then articulated in the expected way as the closure of immediate supervenience: the unordered pair $\{e_1, e_2\}$ *supervenes* on $\{e_3, e_4\}$ if and only if there is a sequence $\langle \{e_1, e_2\}, \dots, \{e_3, e_4\} \rangle$ such that each pair in the sequence (except $\{e_3, e_4\}$) immediately supervenes on its successor in the sequence.

In informal terms, that $\{e_1, e_2\}$ supervenes on $\{e_3, e_4\}$ means that, given the causality relation C , $e_1 \neq e_2$ is *determined* by the fact that $e_3 \neq e_4$. Note that in davidsonian graphs, overdetermination of difference facts is possible.

Aside from difference facts that supervene on other difference facts, there are also difference facts that do not supervene on other difference facts but on causality facts alone. We say that a pair $\{e_1, e_2\}$ is *independent* if and only if

- there is an e_3 such that $e_3 C e_1$ and there is no e such that $e C e_2$, or
- there is an e_3 such that $e_3 C e_2$ and there is no e such that $e C e_1$, or
- there is an e_3 such that $e_1 C e_3$ and there is no e such that $e_2 C e$, or
- there is an e_3 such that $e_2 C e_3$ and there is no e such that $e_1 C e$.

The idea behind this definition is that in order to decide the identity of independent vertices in a graph, no difference facts need to be established.

It is clear that the relation of immediate supervenience and the relation of (mediate) supervenience are both symmetric. This symmetry is due to the symmetry in Davidson's identity criterion. Moreover, it is clear that the supervenience relation is transitive. Therefore, for any $G = \langle V, C \rangle$, the supervenience relation partitions the set of pairs of vertices in V into equivalence classes. Cells of this partition are called *supervenience cycles*.

In terms of the notion of supervenience and the notion of independence, a notion of *predicativity* (or applicability, or epistemic non-circularity) can now be precisely articulated:

Definition 1 (local predicativity) Davidson's criterion is said to be (locally) predicative on $\{e_1, e_2\}$ on G if and only if the supervenience cycle to which $\{e_1, e_2\}$ belongs contains an independent pair of vertices.

When this condition is satisfied for some $\{e_1, e_2\}$ belonging to some davidsonian graph G , we say that the graph is locally predicative on $\{e_1, e_2\}$. Intuitively, predicative pairs of vertices (on a graph) are such that their difference can be reduced to facts which are not themselves difference facts. One might say that their difference is *grounded* in facts which are not themselves difference facts.⁹

Definition 2 (global predicativity) Davidson's criterion is said to be (globally) predicative for G if and only if G is predicative on all pairs $\{e_1, e_2\}$ taken from its domain.

Davidsonian graphs on which Davidson's criterion is globally predicative are called *predicative* graphs. Davidsonian graphs on which Davidson's criterion is globally non-predicative are called *impredicative* graphs. The idea of course is that our notion of impredicativity of davidsonian graphs captures the somewhat vague Quine-Strawson-Cleland notion of impredicativity as applied to Davidson's criterion.

Now it is time to look at a couple of very simple illustrations. As a first example, consider the two-event universe E_1 whose causal structure is given by: $e_1 C e_2$. This is a davidsonian structure: e_1 and e_2 do not have the same causes and effects. It is easy to see that $\{e_1, e_2\}$ is independent on E_1 . To know that $e_1 \neq e_2$, no *identity questions* need to be decided. It is sufficient to know the facts about the causality relation. In particular, once it is known that e_1 has no causes whereas e_2 has a cause, the matter is clinched. So Davidson's criterion is globally applicable on E_1 . As a second example, consider the two-event universe E_2 whose causal structure is given by: $e_1 C e_2, e_2 C e_1$. This too is a davidsonian structure. But for this universe the Quine-Strawson-Lowe accusation of circularity is justified. In order to decide whether $e_1 = e_2$, one must look at the causes and effects of e_1 and e_2 . Consider the causes first. The cause of e_1 is e_2 and the cause of e_2 is e_1 . So we cannot say whether e_1 and e_2 have the same causes before we have decided whether $e_1 = e_2$. For similar reasons we cannot decide whether e_1 and e_2 have the same effects without deciding whether $e_1 = e_2$ holds. Thus $\{e_1, e_2\}$ immediately depends on itself: we have a dependence circle. Davidson's criterion can therefore not be applied to $\{e_1, e_2\}$ on E_2 : it is globally impredicative for E_2 . As a third example, consider a four-element event structure E_3 that is slightly more interesting. Its causal relation is described as: $e_1 C e_2, e_3 C e_4$. E_3 is a Davidson structure. It is left to the reader to verify that $\{e_1, e_2\}$ is predicative on E_3 whereas

⁹ This notion of groundedness is closely related to the groundedness notion that is investigated in [Leitgeb 2005].

$\{e_1, e_3\}$ is impredicative on E_3 . So this event structure is neither globally applicable nor globally inapplicable.

In this way it becomes clear that whether and to which extent Davidson's criterion is circular depends on the causal event structure that it is applied to. It can be shown that as the number n of events tends to infinity, the proportion of n -event structures that are predicative tends to zero. In sum, if there are only finitely many events, then it seems a priori rather likely that Davidson's criterion, as applied to *our* world, is globally inapplicable. In the absence of further constraints, the causal event structure of the world will be overwhelmingly likely to be globally impredicative.

The moral of this is surely that our knowledge about the event structure of the world cannot be *completely* reduced to descriptions of events and statements that one event is a cause or effect of another event. For in practice, we do know of a host of pairs of event descriptions that they denote different events. This must be because we can bring information to bear on such questions that is not reducible to causality information. For instance, temporal information ("event e_1 occurred before event e_2 ") may help us to decide difference questions. But temporal information can only help if the relation of temporal precedence is not itself reducible to causal relations.

Even if we impose more constraints on the causality relation, it appears unlikely that Davidson's criterion can be used as a decision procedure in many circumstances where a dispute arises over the co-extensiveness of two descriptions of events. But if the metaphysical conception of identity criteria is adopted, *that alone* does not constitute sufficient grounds for rejecting Davidson's criterion. In Anscombe's words: "Davidson's criterion of identity for events isn't a criterion in the sense of a 'way of telling'." [Anscombe 1979, p. 230]

As mentioned above, Davidson's criterion can be interpreted as an expression of the causal relational nature of events. But a stricter implementation of this motivation can be discerned. Let us reconsider the (partially applicable) davidsonian structure E_3 . From a *global* ontological viewpoint, one may take this structure to involve an unnecessary duplication: the causal substructures $e_1 C e_2$ and $e_3 C e_4$ ought to be identified. More in general, one might say that in a possible universe two disjoint causal substructures cannot play exactly the same causal role. So if in a putative model of the universe two disjoint substructures play the same causal role, then this model must be *revised* by identifying these two isomorphic substructures.

Let us make this more precise. Call a causal event structure E a *strongly davidsonian* structure if the following holds: There is no

automorphism a such that (1) $a(P)=Q$ and $a(Q)=P$ for some disjoint substructures P, Q of E and (2) a is the identity transformation everywhere else.

The thesis that only strongly davidsonian structures are to be allowed does not have the form of a criterion of identity for events. In accordance with the stance taken in section 2, we do not put too much stock on this: what matters is the ontological implications of the restriction to strongly davidsonian structures. Incidentally, the restriction to strongly davidsonian structures is naturally expressed as a criterion of identity for local causal structures: if there is an automorphism of the universe transforming a local causal structure x into a disjoint local causal structure y , then x and y are the same local causal structure.

The class of strongly davidsonian structures is a proper subclass of the class of davidsonian structures. E_3 , for instance, is a davidsonian structure but not a strongly davidsonian structure. But the class of strongly davidsonian structures is a proper superclass of the class of applicable davidsonian structures. Consider, for instance the structure E_4 , which is given by: $e_1 C e_2 C e_3 C e_1$. This is a strongly davidsonian structure, yet it is globally impredicative.

More variations on Davidson's criterion of identity for events are conceivable. Quine did not favor an ontology of events linked by causal relations. Perhaps causal relations are indeed relics of an all too metaphysical past. But a Davidson-like criterion of identity for events can be formulated purely in terms of statistical correlations between event types.¹⁰ This identity criterion says that an event type x is identical to an event type y if and only if x and y statistically screen each other off from all other event types:

For all event types x, y : $x=y \Leftrightarrow$ for all event types e that are compatible with x and y : $\Pr(x|e) = \Pr(x|y \cdot e)$ and $\Pr(y|e) = \Pr(y|x \cdot e)$

Let us call this *Douven's criterion*. Quine does not accept events or event types in his ontology [Quine 1985]. So we may safely assume that he would also reject this identity criterion out of hand. But in failing to countenance events and event types, his attitude seems to have been anachronistic rather than forward-looking.

As with the insistence on strongly davidsonian structures, it seems that when Douven's criterion is applied *locally*, it does not have a great deal of plausibility. This is so even if the probability involved is given

¹⁰ This was suggested to me by Igor Douven.

an objective interpretation (such as relative frequency). For consider the following scenario. Balls are colliding in front of a mirror. An event and its mirror event screen each other off from all other events. Yet they instantiate distinct event types. But as with Davidson's criterion and its stricter variant, matters are not so clear when we take a global viewpoint. If one takes event types to be *relational by nature* in a specific way, then Douven's criterion might be taken to be correct for theoretical reasons.

The notion of applicability or predicativity as defined in this section can be generalized to identity criteria for other kinds of entities. I spare the reader the formal details. Informally, it should already be clear that the notion of applicability or circularity that is described here yields the right judgement for the principle of extensionality for classes. The axiom of extensionality is globally applicable, for given the Zermelo-Fraenkel axiom of Foundation, the relevant dependence paths always finitely terminate. Also, it is clear that Locke's criterion is globally impredicative in a *trivial* way.

As a last elaborated example, we look at an identity criterion that has been discussed in connection with structuralism in the philosophy of mathematics. Shapiro has defended the position of *Realist Structuralism* about mathematical theories [Shapiro 1997]. Applied to arithmetic, this roughly means that arithmetic is about an *abstract structure*: the structure of the natural numbers. This structure is self-standing, i.e., it is ontologically independent of the manifold of systems that instantiate it. The structure consists of places that themselves possess no internal structure, but whose identity is completely determined by the relations they have to the other places in the structure.

Keränen argues that the Realist Structuralist is *committed* to the following criterion of identity for places in structures:

$$\forall x, y \in S : x = y \leftrightarrow \forall \phi^S (\phi^S(x) \leftrightarrow \phi^S(y)),$$

where ϕ^S ranges over formulae of the signature of the structure *but contains no proper names referring to places in the structure*. Let us call this criterion STR.

Keränen points out that STR fails for a structure S if and only if S admits nontrivial automorphisms [Keränen 2001, p. 319]. Keränen argues that since many mathematical structures *do* admit nontrivial automorphisms, STR is not universally valid. And so, by contraposition, Realist Structuralism is false.

Keränen's ultimate argument for his thesis that the Realist Structuralist must require that ϕ^S 's in the criterion of identity for a

structure are not allowed to contain parameters referring directly to particular places is presented in the following passage [Keränen 2001, p. 325]:

[...] the realist structuralist is supposed to be providing a primitive, indigenous account for the places of any given structure. Her account of identity is supposed to tell us which *particular* identity statements, according to her theory, are true in the language of S . [...] Consider once again the [additive group] structure $(Z, +)$. The formulae $x+x = \underline{2}$ and $x+x = -\underline{2}$ could serve to distinguish 1 from -1 only provided $\underline{2} = -\underline{2}$ is true in the language of $(Z, +)$. However, *this* is precisely what the realist structuralist is not entitled to suppose until she has given an account of identity that entails $\underline{2} = -\underline{2}$.

The problem that Keränen raises is once again an *applicability problem*. But to repeat, this is an *epistemic* matter, whereas Realist Structuralism is an ontological thesis. For this reason, Keränen's argument for disallowing names referring to particular places in the structuralist criterion of identity does not withstand critical scrutiny.

As Keränen himself concedes, when names referring directly to places are admitted, there is no problem for the structuralist criterion of identity. To illustrate this on the basis of a simple example, consider the simple group structure $(\{-1, 0, 1\}, + \text{ mod } 2)$. Let the structuralist identity criterion for places which is just like STR except that names referring to places are allowed in ϕ^S be called STR*. Then STR* entails:

$$-\underline{1} = \underline{1} \Leftrightarrow \forall \phi^S (\phi^S(1) \leftrightarrow \phi^S(-\underline{1}))$$

This yields the right result. Taking $\phi^S(x) \equiv x = 1$ makes the right-hand-side false, whereby $-\underline{1} = \underline{1}$ must be false too. But $\{-\underline{1}, \underline{1}\}$ will not be *predicative* on $(\{-1, 0, 1\}, + \text{ mod } 2)$.

Even though he opposes realist structuralism for reasons that are related to those of Keränen, MacBride arrives at a conclusion that on the face of it appears to resemble ours [MacBride 2006, p. 68]:¹¹

[...] [The realist structuralists] tacitly allow for the possibility that objects are *impredicatively* defined. The numerical diversity of (at least) some objects depends upon their bearing an irreflexive relation to one another. Speaking figuratively [...], there are no diverse objects in advance of the relevant relations obtaining. [...] [The realist structuralist] must conceive of the numerical diversity of

¹¹ For a defense of realist structuralism against recent critiques by MacBride, Keränen and others, see [Leitgeb & Ladyman 2008].

objects and the obtaining of irreflexive relations among them as interdependent (without the dependency between them being viciously circular).

And MacBride calls for attaining a “synoptic understanding of the different forms of impredicativity” that are relevant here. However, a suggestion that emerges from the foregoing is that the familiar notion of *semantic* impredicativity is the wrong concept to focus on in this connection. Rather, the distinctive kind of impredicativity is an epistemological one. Hopefully our discussion has shed some light on it.

7. Summary

The philosophical literature abounds with disputes over the acceptability of proposed identity criteria. Some of these disputes revolve around putative counterexamples to a proposed criterion. Others gravitate around adequacy conditions that satisfactory identity conditions ought to meet. If we are to achieve any form of resolution of these latter discussions, the epistemic component of the traditional concept of identity criterion must be clearly separated from its metaphysical (or, more broadly: theoretical) component.

Once this is done, we are faced with a choice. Do we restrict the honorific criterion of identity to principles that satisfy the epistemic property of (global) applicability, or are we in principle prepared to accept as satisfactory identity criteria some principles that do not meet this requirement? Clearly this threatens to degenerate into a verbal dispute. Nevertheless, there are methodological reasons for refraining from including applicability in the list of adequacy conditions for identity criteria.

If we should agree with Strawson that a principle of identity that is not globally applicable should not be called a criterion, then there may exist substantial kinds of objects for which no identity criteria exist. Perhaps events or places in structures fill this bill. But even for such kinds of objects there may well be natural principles that capture their identity conditions. They deserve scrutiny: we would have to invent a name for them. This, then, seems to me one reason for not building the requirement of applicability into the very concept of criterion of identity.

Another reason is that the concept of applicability or predicativity *itself* merits careful scrutiny. By defining the notion of impredicativity of identity criteria that Strawson and others have objected to, a clearer view is obtained of the ontological implications of their viewpoint. It is a nontrivial question *under which circumstances* an identity criterion or

however one cares to call a principle which features all the hallmarks of an identity criterion except perhaps predicativity, is applicable. And it is a nontrivial task to formulate a general theory of the extent to which identity criteria are applicable. If one opts for the wider concept of identity criterion, such matters are questions of classification for identity criteria.

As always, when all is said and done the proof of the pudding is in the eating. The strongest motivation for taking identity criteria to be metaphysical or purely theoretical principles consists in the philosophical clarifications that this brings about. I hope to have shown how clearly separating questions of admissibility from questions of predicativity can contribute to disentangling disputes that have surrounded uses of identity criteria in various philosophical doctrines. In addition and perhaps more importantly, we have seen in some detail how globally predicative, partially predicative and impredicative identity criteria can function as Ockham's razors, and how the closeness of the shave is determined by the type of razor used.

References

- Anderson, C.A. Analyzing analysis. *Philosophical Studies* 72(1993), p. 199–222.
- Anscombe, E. Under a description. *Noûs* 13(1979), p. 219–233.
- Cleland, C. On the individuation of events. *Synthese* 86(1991), p. 229–254.
- Davidson, D. Identity conditions for events. In: N. Rescher (ed.) *Essays in Honor of Carl Hempel*. Reidel, 1970, p. 216–234.
- Reply to Quine on events. In: [Lepore & McLaughlin 1985, p. 172–176].
- De Clercq, R. A criterion of diachronic identity based on Locke's Principle. *Metaphysica* 6(2005), p. 23–38.
- and Horsten, L. Closer. *Synthese* 146(2005), p. 371–393.
- Frege, G. *The Foundations of Arithmetic. A logico-mathematical enquiry into the concept of number*. Eng. Translation by J. L. Austin, Second Revised Edition, Northwestern University Press, 1980.
- Gödel, K. Russell's mathematical logic. In: Feferman, S. et al (eds.) *Kurt Gödel. Collected Works. Volume II: publications 1938–1974*. Oxford University Press, 1990, p. 119–141.
- Hale, B. and Wright, C. *The Reason's Proper Study. Essays towards a neo-Fregean philosophy of mathematics*. Oxford University Press, 2001.

- Horsten, L. and Leitgeb, H. How abstraction works. In: Hieke, A. and Leitgeb, H. (eds.) *Reduction, Abstraction, Analysis*, Ontos Verlag, 2009, p. 217–226.
- Jubien, M. The myth of identity conditions. *Philosophical Perspectives* 10: Metaphysics, 1996, p. 343–356.
- Keränen, J. The identity problem for realist structuralism. *Philosophia Mathematica* 9(2001), p. 308–330.
- Leitgeb, H. What truth depends on. *Journal of Philosophical Logic* 34(2005), p. 155–192.
- Leitgeb, H. Abstraction grounded. Unpublished manuscript, 2009.
- Leitgeb, H. Criteria of identity, strong and wrong. Unpublished manuscript, 2010.
- Leitgeb, H. and Ladyman, J. Criteria of identity and structuralist ontology. *Philosophia Mathematica* 16(2008), p. 388–396.
- E. Lepore & B. McLaughlin (eds) *Actions and Events. Perspectives on the philosophy of Donald Davidson*. Basil Blackwell, 1985.
- Linnebo, Ø. Predicative fragments of Frege arithmetic. *Bulletin of Symbolic Logic* 10(2004), p. 153–174.
- Locke, J. *An Essay concerning Human Understanding*. Edited with a foreword of Peter H. Nidditch, Oxford University Press, 1975.
- Lowe, J. Substance, identity and time. *Proceedings of the Aristotelian Society Supplementary Volume* 62(1988), p. 61–78.
- Impredicative identity criteria and Davidson’s criterion of event identity. *Analysis* 49(1989a), p. 178–181.
- What is a criterion of identity? *Philosophical Quarterly* 39(1989b), p. 1–21.
- One-level versus two-level identity criteria. *Analysis* 51(1991), p. 192–194.
- *The Possibility of Metaphysics*. Oxford University Press, 1998.
- MacBride, F. What constitutes the numerical diversity of mathematical objects? *Analysis* 66(2006), p. 63–69.
- Merricks, T. There are no criteria of identity over time. *Noûs* 32(1998), p. 106–124.
- Noonan, H. Substance, identity and time. *Proceedings of the Aristotelian Society Supplementary Volume* 62(1988), p. 79–100.
- Quine, W.V.O. Events and reification. In: [Lepore & McLaughlin 1985, p. 162–171].
- Shapiro, S. *Philosophy of Mathematics: structure and ontology*. Oxford University Press, Oxford, 1997.
- Strawson, P.F. [1976] Entity and identity. In: [Strawson 1997, p. 21–51].
- *Entity and Identity and Other Essays*. Clarendon Press, 1997.
- Wiggins, D. *Sameness and Substance*. Basil Blackwell, 1980.

- Williamson, T. Criteria of identity and the Axiom of Choice. *Journal of Philosophy* 83(1986), p. 380–394.
- *Identity and Discrimination*. Basil Blackwell, 1990.
- Fregean directions. *Analysis* 51(1991), p. 194–195.