RAFAEL DE CLERCQ and LEON HORSTEN

CLOSER

ABSTRACT. Criteria of identity should mirror the identity relation in being reflexive, symmetrical, and transitive. However, this logical requirement is only rarely met by the criteria that we are most inclined to propose as candidates. The present paper addresses the question how such obvious candidates are best approximated by means of relations that have all of the aforementioned features, i.e., which are equivalence relations. This question divides into two more basic questions. First, what is to be considered a 'best' approximation. And second, how can these best approximations be found? In answering these questions, we both rely on and constructively criticize ground-breaking work done by Timothy Williamson. Guiding ideas of our approach are that we allow approximations by means of overlapping equivalence-relations, and that closeness of approximation is measured in terms of the number of mistakes made by the approximation when compared to the obvious candidate criterion.

1. INTRODUCTION

According to Williamson, criteria of identity are typically of the form:[1]

$$(1) \qquad \forall x \forall y : f(x) = f(y) \leftrightarrow \Phi(x, y).$$

Here the range of the function f consists of a kind K of objects (persons, perceived colors, meanings, ...) for which a criterion of identity is sought, and the domain consists of the entities in terms of which the criterion of identity needs to be expressed (person-stages, color stimuli, sentences ...). Φ specifies the conditions under which $f(x)$ and $f(y)$ are supposed to be identical.

Since the left-hand-side of (1) is an equivalence relation, the right-hand-side must be an equivalence relation too. For many kinds of objects K, there is an obvious candidate R for playing the role of Φ. This then yields a concrete proposal:

$$\forall x \forall y : f(x) = f(y) \leftrightarrow R(x, y).$$

Unfortunately, such proposals will often be unsatisfactory. As certain long-standing discussions (e.g., concerning personal identity) have made clear,

Synthese (2005) 146: 371–393
DOI 10.1007/s11229-004-6219-1

the obvious candidate R is seldom adequate in all of the aforementioned logical respects. In particular, the relations that have been proposed more often than not fail in respect of transitivity. Still, their initial plausibility may count in favor of them, and so an interesting question is how we can approximate these – at least typically – reflexive and symmetrical relations as closely as possible by relations that do have all of the aforementioned characteristics, that is, by equivalence relations. In other words, the question is how can we obtain logical adequacy while staying as true as possible to the original proposal.

Williamson considers two ways of constructing an adequate substitute relation. One way is to search for a smallest equivalence relation R^+ such that $R \subseteq R^+$.[2] Let us call this the *approach from above*. Such R^+ always exists and is always unique. Another way is to look for a largest equivalence relation R^- such that $R^- \subseteq R$.[3] Such an R^- always exists (on the assumption that the Axiom of Choice holds).[4] but is not typically unique. We will call this the *approach from below.* Williamson notes that one can attempt to mitigate the embarrassment of multiplicity by singling out some R^-s as somehow preferable over others. For instance, one could prefer R^-s with a minimum number of equivalence classes. Williamson calls this the *Minimality Constraint.*[5]

The approach from above and the approach from below then give rise to official candidates for criterion of identity for Ks:

$$\forall x \forall y \, : \, f(x) = f(y) \leftrightarrow R^+(x, y) \, ,$$

$$\forall x \forall y \, : \, f(x) = f(y) \leftrightarrow R^-(x, y) \, .$$

The initial plausibility of the relation R makes it likely that both of these technical substitutes will be unfaithful to (at least some of) our original intuitions concerning identity and difference of K-objects. In particular, the approximation from above will judge some $f(x)$, $f(y)$ to be identical which are really different from each other according to R, whereas the approximations from below will judge some $f(x)$, $f(y)$ to differ from each other which are really identical according to R.

In both Williamson (1986) and Williamson (1990) the two approaches are discussed at length.[6] However, neither work makes mention of the possibility of a third approach, which seeks to construct *partially overlapping* approximations. We suspect that this approach was ignored because of the following reasoning:

If we take a relation ... to be necessary and sufficient for ... identity, and then discover it to be non-transitive, we may give up *either* the necessity *or* the sufficiency. (Williamson 1990, 120; italics not in original)

Giving up the necessity or the sufficiency of the proposed criterion is supposed to define "an obvious pair of fall-back positions" (1986, 382). Evidently, once this is accepted, only two possibilities are left open: the approach from below, in case the proposed criterion is held to be a necessary [but not sufficient] condition for the identity of $f(x)$s, and the approach from above, in case the proposed criterion is held to be a sufficient [but not necessary] condition for the identity of $f(x)$s.

However, to say that these fall-back positions are somehow obvious is not to say that they exhaust the field of possibilities. Indeed, a third possibility suggests itself, namely that of regarding the proposed criterion as *neither* necessary *nor* sufficient for the identity of $f(x)$s. At least in Williamson (1986) and Williamson (1990), Williamson offers nothing to preclude this possibility, which leads naturally to the approach we want to consider and defend in this paper. Moreover, it seems plausible to assume that if we can be mistaken in taking a criterion to be both necessary and sufficient – as Williamson seems ready to admit – we can also be mistaken in taking it to be either of these. Yet it seems that even in such cases we need not be far from the truth in thinking about the criterion as a plausible candidate.

Admittedly, there are cases where a proposed criterion is without doubt necessary for identity. For instance, being perceptually indistinguishable is a plausible candidate criterion of identity for perceived color, and is no doubt a necessary condition for it. Similarly, certain forms of mental continuity may be regarded as sufficient conditions for personal identity. But this is perhaps already less obvious. More generally, it seems that obviousness comes in degrees, and probably in different degrees to different persons. And as long as there is no compelling reason to regard a condition either as obviously necessary or as obviously sufficient (as will often be the case) it seems more reasonable to keep all options open instead of retreating immediately to either of the two fall-back positions considered by Williamson.

In what follows, an entirely general approach to approximating relations is described. We will call it the *overlapping approach*. In our view, this approach is superior to Williamson's not just because it is able to generate the right kind of approximation in cases where the (approximated) criterion turns out be neither a necessary nor a sufficient condition for identity. It is also superior because, as will be shown, it is able to generate *closer* approximations. In particular, the idea is that equivalence relations which partially overlap R are often able to more closely approximate R than sub- or super-relations of R that are equivalence relations.[7]

Of course, our approach leads to closer approximations only relative to a given standard of closeness. Here too we would like to make a contribution to Williamson's project. More specifically, we think that the standard of closeness proposed by him can be improved in respect of precision. Furthermore, we think that this increase in precision can be maintained in the infinite case – contrary to what Williamson seems to have feared.

In what follows, we will make the idealizing assumption that the candidate criterion R is *discrete*: two objects either clearly stand in the relation R to each other or they clearly do not. This assumption is made merely to avoid unnecessary complexities in the presentation of our approach. Nonetheless, we will briefly return to it in the final section of the paper, where it is described how our approach can be refined to suit more realistic scenarios.

Finally, we would like to make a brief comment on the intended purpose of this paper. We do not claim to be solving any specific philosophical problem here. The improvements we propose are mainly of a methodological or meta-philosophical kind. Nonetheless, it should be kept in mind that, since Williamson's approach is clearly a variant or subspecies of ours, we are at least able to deal with the problems tackled by him in the aforementioned works, for instance, all sorts of sorites paradoxes involving the concept of identity (more specifically, identity of phenomenal character, transworld identity, and identity through time). Moreover, the applicability of our approach is of course not restricted to cases where the concept of identity is involved. Our approach applies more generally to all cases where an approximating equivalence relation is to be constructed on the basis of a given reflexive symmetrical relation. Finally, it may be noteworthy that non-philosophical applications have already been found by authors working in applied mathematics.[8]

2. FOUNDATIONAL ISSUES

To make sure that our project makes sense from a philosophical point of view, we would like to address some questions concerning its foundation. For instance, it might be asked what the search for a technical substitute is supposed to achieve. Clearly, the technical substitute is supposed to replace the obvious candidate, which for logical reasons we cannot hold on to. But this replacement can still be interpreted in two ways, depending on what the status of the obvious candidate is supposed to be: the obvious candidate may be regarded either as a first articulation of our common sense criterion of identity, or as the common sense criterion itself. Accordingly, the search for a technical substitute may be regarded either as an attempt to *make* *explicit* what organizes our practice of identification, or as an attempt to *improve* – at least from a logical point of view – this practice. In this paper, we do not arbitrate between these two interpretations. We merely wish to note that the choice is likely to depend, firstly, on the particular criterion of identity at issue, and secondly, on the kind of shortcomings one is prepared to ascribe to common sense in general, and to our practices of identification in particular.

A related question is whether the methods we describe are to be seen as part of a reductionist strategy. After all, the aim seems to be to construe objects as equivalence classes of more basic entities, and this is certainly remindful of certain reductionist programmes belonging to the early days of analytic philosophy (in particular, of Carnap's *Aufbau*).[9] However, although our methods may well be fitted into such programmes, they are not in themselves of a reductionist sort. Firstly, by stating the identity conditions of an entity of kind A in terms of relations holding between entities of kind B, we do not automatically identify entities of kind A with collections of entities of kind B. Secondly, even if this kind of identification were legitimate (for one reason or another), applying the methods described in this paper would still not lead to large-scale reduction without the assumption that there is some kind of basic entity (say, elementary experiences) out of which all the others can be constructed as equivalence classes, or as equivalence classes of equivalence classes, etc.[10] But for all we know our methods are only sporadically applicable (to a gerrymandered set of entities) and not systematically until one reaches the lowest layer of reality.

Yet another interesting question is whether we should search for a 'meaningful' or definable technical substitute.[11] One reason why we should do so is this. If we cannot attach any meaning (or sense) to the newly concocted identity criterion, i.e., if we cannot grasp it in any other way than by enumerating the elements of its potentially infinite extension, then we are unlikely to articulate something that is already present in the practice of finite minds such as our own. Moreover, we are unlikely to propose something that might once improve our practice by becoming a part of it. In other words, whether our goals are revisionary or merely explicatory, a definable relation seems to be an uncontroversial desideratum, at least as long as the identity criterion is supposed to play a role in our practice. However, this is not to suggest that the requirement of 'meaningfulness' is to be regarded as absolute. Other considerations such as 'closeness of approximation' may well outweigh the meaningfulness condition in certain circumstances. Moreover, as far as we can see, there is no guarantee that our conceptual resources will always suffice to define any of the best approximations to a given relation. Finally, the existence of a fast algorithm for generating best approximations to a particular candidate criterion may also be regarded as a way of satisfying the Meaningfulness Constraint. After all, if no fast procedure for doing this exists, then people cannot reliably find adequate technical substitutes in a reasonable amount of time.

If ordinary criteria of identity are considered open to improvement, then the cases studied by Williamson can be seen as instances of a more general phenomenon, namely the 'imperfection' – from a theoretical or formal point of view – of ordinary language and its associated concepts. This imperfection is manifested in many ways, and it is also reflected in the area that interests us here. Let us explain this a bit further.

Ordinary language concepts are imprecise and they usually resist analysis into simpler concepts that are better understood. Moreover, their application is seldom governed by a set of clearly defined rules. Nonetheless, these concepts serve our purposes, and where needed, a precisification or formal counterpart is usually not too hard to get. Similarly, it seems that many of the things falling under these concepts lack identity criteria that are theoretically wholly adequate. For instance, for most identity criteria there exist marginal cases that cannot be decided by them, and in certain (actual or merely imagined) circumstances the transitivity of a criterion may break down. However, these shortcomings need not impede us from applying a certain criterion in ordinary, day-to-day circumstances. After all, the actual state of the world is such that usually a less adequate criterion will do. Moreover, a theoretically more adequate criterion may be hard to find, difficult to articulate, or cumbersome in its application. Thus, it should not be found surprising that many of the criteria we use in deciding issues of identity are in fact theoretically inadequate. They are merely as adequate as they need to be, just as our concepts are as clear and precise as they need to be (given our purposes and the actual state of the world). It is only the philosopher, or the scientist, who is inclined to complain about them.

3. OVERLAPPING APPROXIMATIONS

In this section, a way of constructing equivalence relations is described which yields closer approximations to the intuitive non-equivalence relation R than R^+ and the R^-s from the first section. As mentioned earlier, the underlying idea is that equivalence relations which partially overlap R are often able to more closely approximate R than sub- or super-relations of R that are equivalence relations.[12] The idea can best be illustrated with a simple concrete example.

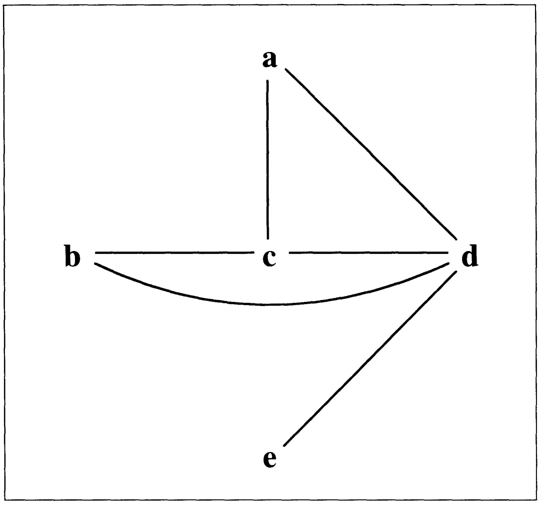

Figure 1.

Consider the following underlying domain of objects:

$$D = \{a, b, c, d, e\}\,.$$

The underlying common sense relation R is assumed to be reflexive and symmetrical. When a relation holds between (distinct) objects x and y, we denote this as $\overline{xy}$. The relation R can then be uniquely specified as a collection of elements of the form $\overline{xy}$. The following will be our definition of the intuitive relation R:

$$R = \left\{\overline{ac}, \overline{ad}, \overline{bc}, \overline{bd}, \overline{cd}, \overline{ed}\right\}.$$

Alternatively, R can be represented as a simple undirected graph (Figure 1). Again, it is assumed that R is reflexive and symmetrical. But this particular R is obviously not an equivalence relation.

It is easy to see that R^+ is the universal relation on D. To obtain R^+, we just have to 'complete all the open triangles' in Figure 1.

The following relation is one of the maximal equivalence relations Iinclude in R:

$$R^- = \left\{\overline{bc}, \overline{bd}, \overline{cd}\right\}.$$

R^- can also be represented as a graph (Figure 2).

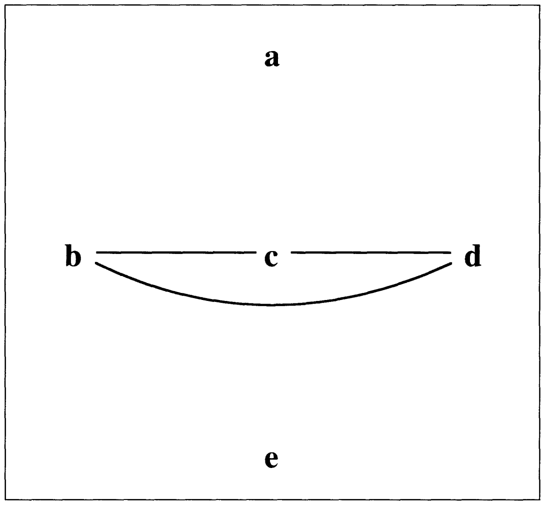

Figure 2.

The *degree of unfaithfulness* of R^+, R^- relative to R can be measured by counting the number of revisions necessary to get R^+, R^- from R. For instance, to obtain R^+ we have to add 4 edges to R, which yields an unfaithfulness degree relative to R of $|+4| = 4$. Similarly, obtaining R^- requires removing 3 edges from R, which means that its unfaithfulness degree will be $|-3| = 3$. As a result, R^- is slightly better than R^+, for it commits one less violation against our intuitions as expressed by R.

Now consider the following equivalence relation, which partially overlaps R:

$$R^{\pm} = \left\{\overline{ab}, \overline{ac}, \overline{ad}, \overline{bc}, \overline{bd}, \overline{cd}\right\}.$$

Graphically, $R^{\pm}$ can be presented as in Figure 3.

When compared with R, we see that $R^{\pm}$ adds one edge ($\overline{ab}$) and removes one ($\overline{ed}$). So the degree of unfaithfulness of $R^{\pm}$ is $1 + 1 = 2$. This means that $R^{\pm}$ is less unfaithful than R^-, i.e., $R^{\pm}$ respects our intuitions even better than R^-. In this precise sense, overlapping equivalence relations can provide closer approximations to intuitive (but unacceptable) criteria of identity than approximations from below or from above. What is more, the advantage of overlapping over the approach from below and the approach from above can be shown to be unbounded. Replacing the clique $\overline{bcd}$ of R by a clique $\overline{b_1 \ldots b_n}$ which is almost completely connected to a and almost completely unconnected to e yields for sufficiently large n an arbitrarily large advantage of overlapping over approaching strictly

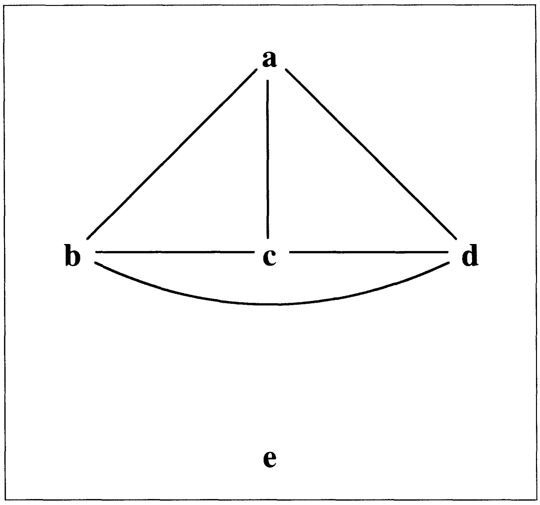

Figure 3.

from below or from above. Suppose, for instance, that there is only one connection between $\overline{b_1 \ldots b_n}$ and e, while there is only one connection lacking between $\overline{b_1 \ldots b_n}$ and a (say, the connection between b_1 and a.) The approach from below will then have to cut all connections between $\overline{b_1 \ldots b_n}$ and a. (Alternatively, i.e., making an equal number of mistakes, it may cut all connections between b_1 and $\overline{b_2 \ldots b_n}$.) In addition it will have to cut a connection with e. Conversely, the approach from above will have to make $\overline{b_1 \ldots b_n}$ and e entirely connected. Moreover, it will have to add an extra connection with a. The overlapping approach, however, only requires one connection to be added and one to be omitted.

4. RANKING APPROXIMATIONS

Williamson (1986) briefly considers ranking *approaches from below* according to the number of mistakes they make. Williamson then rejects this option for the following reason:

> One might look for finer-grained considerations that would single out some of the relations that [are maximal approaches from below] from the others. For example, one might try to minimize the *cardinality* of that part of the extension of [the approximation] which is disjoint from the extension of the criterion of identity at issue. However, the source of the divergence between criteria of identity is essentially symmetrical – two things related to a third thing but not to each other – so there is reason to think that non-arbitrary finer-grained considerations will make little difference. (1986, 389)

In Williamson (1990), the objection seems to be that counting mistakes would lead to "complications" (p. 123).[13] Williamson therefore prefers a "crude but simple" standard of closeness. But by calling it a "crude" standard, and by elsewhere conceiving of best approximations in terms of "minimal revisions" (1990, 109), and "minimum mutilation" (pp. 115, 123) he suggests that a more precise standard in terms of counting pairs would be appropriate, although probably unpractical. In the following subsections, we will show why this worry was unfounded. Sections 4.1 and 4.3 present a purely quantitative standard of closeness which avoids the complications alluded to by Williamson. Section 4.2 presents his own 'qualitative' standard.

4.1. *A Quantitative Approach*

In Section 3 we already employed a quantitative standard of closeness by comparing degrees of unfaithfulness. In this section our aim is to render this standard fully explicit in a series of definitional steps.

DEFINITION 1. For any given finite R, and any equivalence relation E on dom(R), let $\mu_R^+(E)$ be the set of edges belonging to E but not to R, and let $\mu_R^-(E)$ be the set of edges belonging to R but not to E.

DEFINITION 2. For R and E as in the previous definition, let $\mu_R(E)$ (the set of mistakes of E) be defined as

$$\mu_R^+(E) \cup \mu_R^-(E)\,.$$

When the context permits it, we will in the sequel often omit the subscript R.

In words, the idea can be expressed as follows. An equivalence-approximation E is obtained from R by removing and adding edges to/from R (cutting and pasting). Every action of adding and of omitting an edge is counted as a 'mistake' made by E.

DEFINITION 3. For R and E as before, E is a quantitatively best approximation to R if and only if the size of $\mu_R(E)$ is minimal.

In other words, a quantitatively best approximation to R is an approximation which makes a minimal number of mistakes. This definition goes back at least to Zahn (1964, 840). It defines what is quantitatively best because the mistakes are *counted*.

The notions of best approximation "from above" and "from below" can now be defined in terms of the previous, more general definition: *the* quantitatively best approximation from above to R is the unique quantitatively best approximation E (in the sense of Definition 3) with $\mu_R^-(E) = \emptyset$, and *a* quantitatively best approximation from below to R is a quantitatively best approximation E (again, in the sense specified by Definition 3) with $\mu_R^+(E) = \emptyset$. In other words, best approximations from above are best approximations that only add edges to R (but do not omit edges), whereas best approximations from below are best approximations which merely remove edges from R (but do not add any).

4.2. *A Qualitative Approach*

Williamson's notions of best approximation from below (R^-) and best approximation from above (R^+) may be called *qualitative* because they do not require one to count the number of revisions necessary to obtain an approximation by modifying a given obvious candidate R. In fact, these notions can be regarded as special cases of an overarching qualitative notion of best cut-and-paste approximation which can be expressed as follows:

DEFINITION 4. E is a qualitatively best approximation to R if and only if E is an equivalence relation on $\text{dom}(R)$ and there is no equivalence relation E' such that $\mu_R\left(E'\right) \subset \mu_R(E)$.

In other words, E is a qualitatively best approximation of R if and only if further progress (relative to E) can only be made by also at some places going against the original graph R. Of course, if E is a best equivalence-approximation in the quantitative sense, then E is also a best equivalence-approximation in the qualitative sense. However, the reverse is not true. For example, the approximation represented by Figure 2 is best *only* in the qualitative sense of the word.

The qualitatively best approximation from above to R (R^+) can then be defined as the unique qualitatively best approximation E (in the sense of Definition 6) with $\mu_R^-(E) = \emptyset$. Similarly, a qualitatively best approximation from below to R (R^-) can be defined as qualitatively best approximation E (again, in the sense specified by Definition 6) with $\mu_R^+(E) = \emptyset$.

4.3. *Infinite Graphs*

One advantage of the qualitative notion of a best equivalence-approximation described above is that it is immediately applicable to infinite graphs. Our quantitative notion of best equivalence-approximation, in contrast, cannot be applied to infinite graphs without further ado. Two (denumerably) infinite equivalence approximations may both make infi-

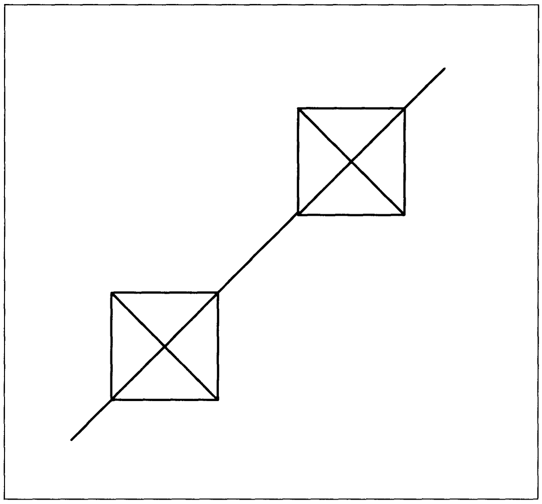

Figure 4.

nitely many mistakes, and yet one might intuitively be thought closer to R than the other. Imagine, for example, that the chain-like structure of Figure 4 stretches out into infinity, always repeating the same pattern. Then consider the following two equivalence-approximations to the relation represented by this structure. One approximation cuts all links between distinct elements, the other merely cuts links between subsequent diamonds. If the quantitative approach were naively applied to this case, the two approximations would be reckoned equally good because they both make an infinite number of mistakes. Nonetheless it is intuitively clear that the second approximation should be preferred. Somehow it seems less mutilating.

In what follows, we want to express a quantitative notion of better equivalence-approximation which accounts for this intuitive difference and which in a sense reduces to the situation for finite graphs, where we can count mistakes. We restrict ourselves to graphs in which all vertices have finite valence; we call such graphs – even if they have an infinite set of edges – *finitary*. The idea is that an equivalence relation E_2 is at least as close to a denumerably infinite but finitary graph R as E_1 if and only if E_2 is on every sufficiently large *local scale* at least as close to R as E_1 is. This would entail some sort of reduction of the notion of better equivalence-approximation to the finite.

DEFINITION 5. Let S be any finite subset of $dom(R)$. Then the 1-*step extension* of S is defined as:

$$\text{ext}(S) \equiv S \cup \{a \mid \exists b \in S : ab \in R\}.$$

DEFINITION 6. For S as before, let $\text{ext}^m(S)$ be $\underbrace{\text{ext}(\ldots \text{ext}(S))}_{m \text{ times}}$.

DEFINITION 7. Let $R = (V, E)$ be given. Then we say that $E_1 \preceq E_2 \Leftrightarrow$ for every finite subset $S \subseteq V$, there is an n such that for all $m \geq n$:

$$\mu(E_1 \upharpoonright \text{ext}^m(S)) \leqslant \mu(E_2 \upharpoonright \text{ext}^m(S)).$$

In words, every finite subset of the original relation can be extended an infinite number of times following the principles outlined in Definitions 5 and 6. Very roughly, Definition 7 states that an approximation E_1 is *at least as good a*s an approximation E_2 if and only if along a series of subsequent extensions of a finite subset S of R there is a point from which E_1 becomes at least as good as E_2 according to the definition for *finite* sets. (Note by the way that for this definition to make sense it has to be assumed that R is finitary.) An approximation E_1 is then quantitatively *better* than an approximation E_2 if and only if E_1 is at least as good as E_2 but not conversely. Obviously, an approximation E is quantitatively *best* if and only if there is no quantitatively better one.

The idea behind this definition is the following. We want to allow for the possibility that E_1 is a worse approximation than E_2 even though on some (or indeed many) finite subsets S of R, E_1 fares quantitatively better than E_2. However, when any such S is expanded along R to a sufficiently large (but finite) scale, E_2 must at some point start to be quantitatively superior to E_1 and from thereon continue to be so.

It is clear that for finite R, these definitions reduce to the quantitative notions of better/best equivalence-approximations that were formulated in Section 2. Moreover, they too imply that quantitatively best approximations are always qualitatively best.

In the Appendix it is demonstrated that not every relation has a best approximation in the sense just defined. In other words, our generalized quantitative definition will not always pick out a best approximation: in some cases there will be an infinite series of ever better approximations. This result may be taken to reveal a flaw internal to our definition of best approximation. But then it is not so clear why every relation should have a quantitatively best approximation. A safer conclusion to draw is that for practical purposes, i.e., in cases where we really need a best approximation, our quantitative definition may have to be backed by another definition such as the qualitative definition proposed by Williamson.

There is room for discussion about the philosophical relevance of infinite graphs. On the one hand, it is a fact that most of the relations we are interested in have a potentially infinite extension (e.g., the relations that underlie sorites paradoxes). On the other hand, what makes these relations interesting from our point of view (failure of transitivity) can be brought out by finite graphs. Thus, nothing of essence seems to be lost when we use such graphs to represent the kind of problems we are faced with, and the kind of solutions we propose. This observation is further supported by the fact that our definition of best approximation in a sense reduces the infinite case to the finite case.

5. REFINEMENTS

The approach from above is the only one which is guaranteed to lead to a unique outcome. However, this approach has as its chief disadvantage that it underwrites sorites reasoning. For instance, according to the official criterion of identity generated by the approach from above, orange and yellow would be the same perceived color; worse even, there would be only one color. In addition, the approach from above will not always lead to approximations that are closest. The same is true of the approach from below. The overlapping approach, however, being the most general one, is guaranteed to lead to approximations that are closest. But this approach shares its main disadvantage with the approach from below: in general, best approximations are not unique.

Of course, where there is more than one best approximation extra conditions can be imposed. For instance, Williamson believes that imposing the Minimality Constraint may be a reasonable thing to do in some cases: "The intuitive effect of the Minimality Constraint is often to sift the sensible maximal M-relations from the silly ones" (1990, 72–73). In the general case, however, Williamson does not recommend the imposition of this constraint (p. 77). And indeed, it seems advisable only where we are really free to furnish our ontology in the most economic way possible. (The Minimality Constraint may be seen as an embodiment of Ockham's Razor because each equivalence cell corresponds to an entity of the kind for which a criterion was sought.)

However, even some of the principles underlying scientific taxonomies seem to offer us this kind of freedom. For instance, two animals may be said to belong to the same biological kind if and only if they are able to

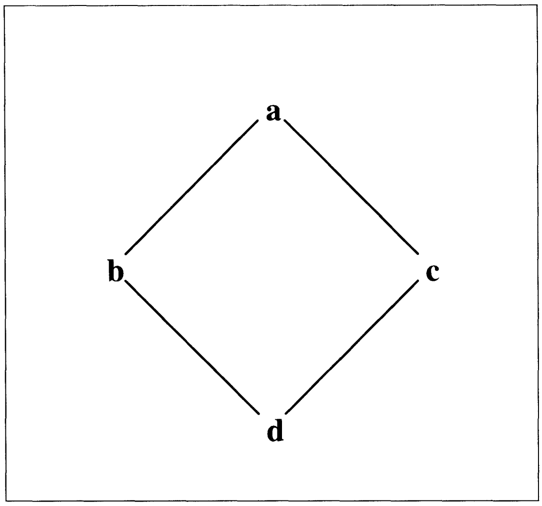

Figure 5.

interbreed or if and only if they exhibit a high degree of morphological similarity. Whichever criterion is chosen, transitivity will break down at some point.[14] Thus approximations will have to be constructed according to one of the principles described above. But here is where problems begin. On the one hand, the approach from above will not deliver a genuine subdivision of kinds since it will group together all animals that are linked directly or indirectly by the criterion of interbreeding or morphological similarity. On the other hand, the approach from below and the overlapping approach will leave us with several subdivisions to choose from. However, here the Minimality Constraint may come to the rescue. After all, simplicity is a plausible constraint on taxonomies.

Note that overlapping approximations are likely to be favored by the Minimality Constraint, since they usually involve fewer cuts than approximations from below. However, note also that opting for a quantitative approach to closeness does not mean that the Minimality Constraint is automatically satisfied. Consider, for example, the relations represented by Figures 5, 6 and 7. Figures 6 and 7 both represent quantitatively best approximations to the relation represented by Figure 5. However, the approximation represented by Figure 7 contains fewer equivalence cells than the approximation represented by Figure 6 (namely one instead of two). In other words, this example shows that a minimal number of mistakes does not imply a minimal number of cells.

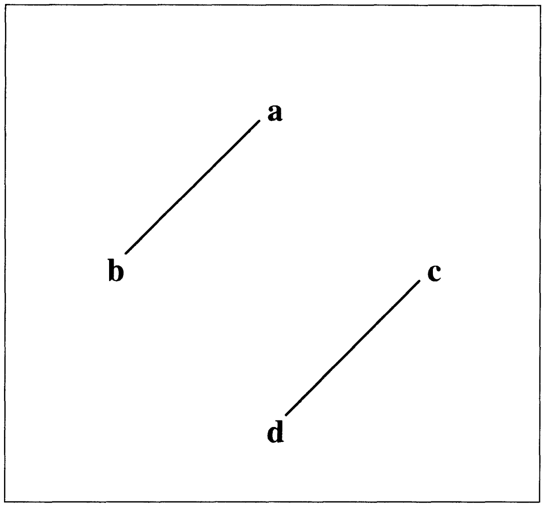

Figure 6.

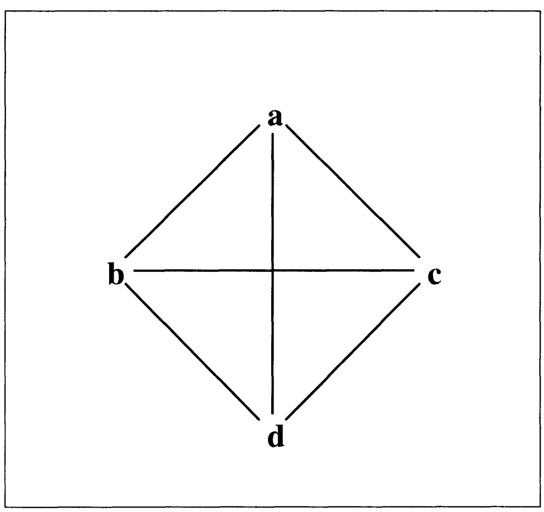

Figure 7.

No doubt there are still more constraints that can be imposed to select among approximations that are quantitatively and/or qualitatively best. *Conservativeness* might be one. Suppose, for instance, that we want to expand the domain of a relation for which we have already found a best approximation. For pragmatic reasons we may well want to stick to this approximation. That is, as we expand the original domain we may be prepared to include new elements in the existing equivalence classes of the approximation; we may even be prepared to create new equivalence classes; but we may not want to effect an entire re-partitioning of the domain by reallocating elements. In the case of biological taxonomies this could be a plausible constraint. After all, we usually want our biological taxa to retain their meaning, to remain applicable to the same organisms, instead of changing their meaning each time the domain of classified organisms is expanded. (The reader is free to check for herself that the most secure way of meeting the requirement of Conservativeness is to combine the approach from below with a qualitative approach to ranking approximations.)

Nonetheless, the imposition of such extra requirements may still leave us with an awkward plurality of best approximations. Moreover, even if there were just one approximation left, the question could still arise whether this is the correct or *true* criterion of identity for the relevant entities.

In sum, the problem of non-uniqueness is a serious one, and also, to some extent, a philosophical one. In this connection Williamson hesitates between two views.[15] On the one hand, there is the view that among the plurality of best approximations there is exactly one correct criterion of identity, although we may never be able to tell which one it is. This is called the *ignorance view*. On the other hand, there is the view that there is no determinate matter of fact with respect to which among a plurality of best approximations is the correct one.[16] This can naturally be expressed using the idea of supervaluation valuations (Williamson 1990, 77), so we will call this the *supervaluation view*. On this view, the final criterion of identity is an *indeterminate* relation R_s, which has an *extension* $\mathcal{E}(R_s)$ and an *anti-extension* $\mathcal{A}(R_s)$ which are defined as follows:

- $\langle a, b\rangle \in \mathcal{E}(R_s) \Leftrightarrow$ for all $E \in BA(R) : \langle a, b\rangle \in E$;
- $\langle a, b\rangle \in \mathcal{A}(R_s) \Leftrightarrow$ for all $E \in BA(R) : \langle a, b\rangle \notin E$.

In general, $\mathcal{E}(R_s) \cup \mathcal{A}(R_s)$ will not exhaust $(\mathrm{dom}(R))^2$, i.e., relying on R_s would often not help us to determine whether two objects a, b belong to the same kind. In such cases R_s could be considered an indeterminate relation. Nevertheless, R_s will always be a (unique) *partial* equivalence relation, i.e., for every R_s, for every a, b belonging to the domain of R: $\langle a, a\rangle \in R_s$; $\langle a, b\rangle \in R_s \Rightarrow \langle b, a\rangle \in R_s$; $\langle a, b\rangle, \langle b, c\rangle \in R_s \Rightarrow \langle a, c\rangle \in R_s$.

However, we should not immediately feel compelled to choose between the two views considered by Williamson. Before making such a choice we could try to soften the problem of non-uniqueness by assigning *weights* to the ordered pairs that belong to the domain of the obvious candidate R. This would also be a way of lifting the simplifying discreteness assumption that was made in Section 1. We could label all $\langle a, b\rangle \in \mathrm{dom}(R)$ with weights $w\left(\overline{ab}\right) \in [0, 1]$. The weight assigned to a pair would represent our degree of confidence in it. In other words, it would indicate the extent to which we are certain that two elements are R-connected: 1 means absolute certainty that they are R-related; 0 means absolute certainty that they are not R-related. Then we could construct equivalence relations to which we apply our measure of unfaithfulness in the same way as before, except that we would now be calculating with real numbers. For instance, if $w\left(\overline{ab}\right) = 0.64$, and according to equivalence-approximation E, a and b are connected, then this decision will increase the unfaithfulness of E with 0.36. As before, a best approximation is one with minimal unfaithfulness. However, it would be a striking coincidence now if two or more approximations still came out as having the same degree of unfaithfulness.

Finally, note that non-uniqueness is not necessarily a disadvantage. It may simply reflect the indeterminacy inherent in our concept of K-identity, and moreover, may enable us to choose that criterion which fits our purposes best in a given context.

6. A LAST LOOK AT RANKING

Approximations can be ranked according to different criteria, e.g., number of equivalence classes (the Minimality Constraint), conservativeness, meaningfulness, and closeness. In this paper we have been chiefly concerned with the criterion of closeness. For instance, one of our central claims is that overlapping approximations are guaranteed to be closest. However, it was also observed that closeness is determined relative to a standard of closeness, and it turned out that there are at least two types of standard: a quantitative and qualitative one. The standard we relied upon in making our claim about overlapping approximations was a quantitative standard. But why should this type of standard be preferred? In what follows, our aim is to compare the respective merits of the two types of standard. (Most of these merits have already been signalled in the preceding sections.) It will emerge that the basic recommendation made in this paper – search for the overlapping approximation which is quantitatively closest – is not to be followed unconditionally, and moreover, that the choice of a particular standard of closeness has definite implications with respect to meeting the other aforementioned criteria.

One advantage of the *quantitative* approach is that it allows us to make very precise judgements of closeness. In particular, it allows us to differentiate between approximations which the qualitative approach is bound to place on a par. Moreover, these further differentiations have an intuitive appeal. Secondly, in general quantitatively best approximations involve commitment to fewer entities because they tend to contain fewer equivalence classes. Thirdly, the non-uniqueness problem discussed in the preceding section is less pressing when a quantitative approach is adopted. After all, every quantitatively best approximation is qualitatively best, but not vice versa. Moreover, when weights are assigned to the pairs in the original relation the non-uniqueness problem virtually disappears altogether.

One important advantage of the *qualitative* approach is that its concepts of best approximation can be carried over from the finite to the infinite case without further ado. Providing a correct definition of quantitatively best approximation to an infinite graph turned out to be more complicated. Moreover, it turned out that not every relation with an infinite domain *has* a quantitatively best approximation.

In fact, there is even more to be said for the qualitative approach. Firstly, a qualitatively best approximation is more likely to be 'preserved' when the domain of a relation is extended.[17] In other words, such an approximation is more likely to meet the Conservativeness Constraint (cf. the previous section). Secondly, and not mentioned earlier, there exist 'fast', i.e., polynomial time, algorithms generating approximations from below that are best in a qualitative sense. In contrast, the problem of finding quantitatively best approximations is NP-complete, i.e., computationally intractable (Krivanek and Moravek 1986; Delvaux and Horsten 2002). It is therefore unthinkable that this problem could be the one human beings try to solve when they attempt to find an adequate criterion of identity. However, this last advantage of the qualitative approach should be qualified in two ways. First of all, it is being presupposed that the approximations are not required to meet the Minimality Constraint, since by imposing this constraint the problem (of finding a qualitatively best approximation) becomes also NP-complete. Secondly, there do exist relatively simple efficient algorithms that generate quantitatively best approximations in *most* cases.

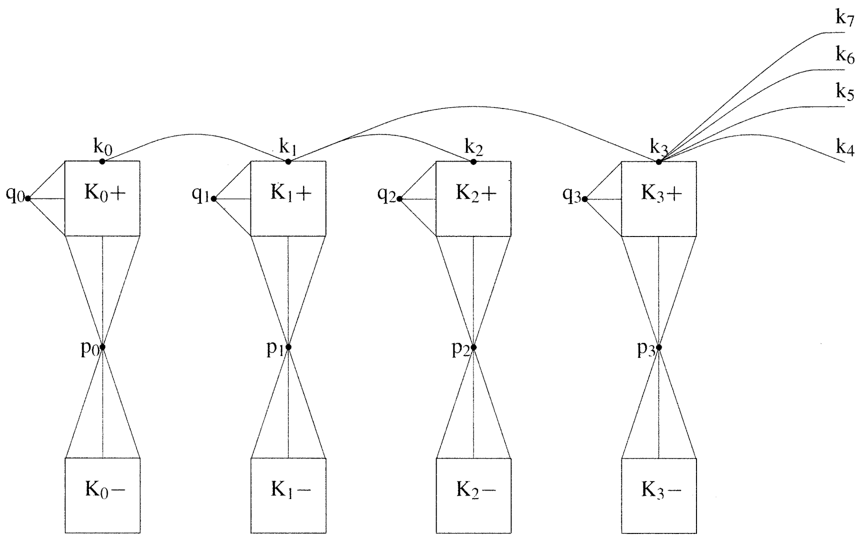

Figure 8.

APPENDIX

In this section we show how to construct an infinite graph that does not have a quantitatively best equivalence-approximation.[18] This graph K is graphically represented by Figure 8.

K consists of components K_0, K_1, K_2, ... which are constructed as follows. Every K_i (with $i \in \mathbb{N}$) contains two cliques K_i^+ and K_i^- which are of equal size. It does not really matter how many points each of these cliques has; for definiteness, say 10. Also, every K_i contains a point p_i that is connected with every point of K_i^+ and K_i^- and with nothing else. Furthermore, K_i contains a point q_i that is connected with every point of K_i^+ and with nothing else. Let k_i be some arbitrary point of K_i^+. Then the components of K are connected to each other in the following way: k_0 is connected to k_1; k_1 is connected to k_2 and k_3; k_3 is connected to k_4, k_5, k_6, k_7; ... k_n is connected to $k_{n+1}, \ldots, k_{2n+1}$.

Consider, first, the finite set F_0, which we set equal to $K_0^+ \cup \{p_0, q_0, k_1\}$. Then look at subsequent 1-expansions F_1, F_2, F_3, ... We see, for instance, that $F_1 = K_0 \cup K_1^+ \cup \{p_1\}$, and in general, an F_n will consist of a number of K_is plus a "tail" of $K_j^+ \cup \{p_j, q_j\}$s plus a number of k_ls. Note that we do not lose any generality by starting with F_0 and comparing how equivalence-approximations fare when restricted to the F_ns.

Next, consider the equivalence-approximations of K. It is easy to see that in every best equivalence-approximation there might be, the components K_i of K are all disjoint, and every component K_i is itself split into 2 disjoint parts. Moreover, every q_i is joined with K_i^+. The only significant question concerns the points p_i: will p_i be joined with $K_i^+ \cup \{q_i\}$ or will p_i be joined with K_i^-? So every non-improvable equivalence-approximation E can be written as an infinite sequence $(+, -, +, +, -, \ldots)$, where at place i we write a $+$ if p_i is joined by E with $K_i^+ \cup \{q_i\}$, and $-$ if p_i is joined with K_i^-.

Note that every E which differs from the 'purely negative' equivalence relation $(-, -, -, -, -, \ldots)$ can be improved by changing its leftmost $+$ into $-$. The reason is this. Let the leftmost $+$ occur at place i. As soon as an expansion F_n is reached which contains the whole of K_i, it is better to change the $+$ into a $-$. And this will remain so in all further expansions. But on the other hand, the purely negative equivalence relation is even worse than the purely positive equivalence relation $(+, +, +, +, +, \ldots)$! The reason is that for every F_n, the advantage of connecting every p_i in the tail of F_n to K_i^+ outweighs the advantage of connecting every p_i in the "main part" of F_n to K_i^-. Therefore K does not have a quantitatively best equivalence-approximation.

ACKNOWLEDGEMENTS

Thanks to Mark van Atten, Lieven Decock, Steven Delvaux, Roland Hinnion, Piotr Kulicki and Timothy Williamson for helpful comments. Rafael De Clercq's research was supported by the Fund for Scientific Research, Flanders.

NOTES

[1] Williamson calls these "second-level" criteria (1990, Chapter 9, Section 9.1). We will here ignore the relation between first-level and second-level criteria because the logical requirements arise anyhow, that is, irrespective of whether second-level criteria are regarded as reducible to first-level criteria.

[2] In Williamson (1986) such an R^+ is called a Maximal A-Approximation. Grice (1941) was probably the first to apply this technique to a philosophical problem. However, see Williamson (1990, 122), for an important note concerning the interpretation of the result.

[3] In Williamson (1986), such an R^- is called a Maximal B-Approximation.

[4] See Williamson (1990, 154–157).

[5] This constraint is introduced in Williamson (1990, 72–73). We will return to it in Section 5.

[6] Because of the particular example on which Williamson concentrates in Williamson (1990) (identity of phenomenal character), most of the book is devoted to the approach from below, resulting in the search for so-called "maximal M-relations". Williamson (1986) is also mainly concerned with the approach from below. However, in Chapter 7 of Williamson (1990), the approach from above is also discussed.

[7] An extra advantage of the overlapping approach might be that it generates approximations with fewer equivalence classes. For more on this feature, see Section 5.

[8] See, for instance, Zahn (1964, 846–847).

[9] Also, the term 'rational reconstruction' seems to apply very well to our approach.

[10] Note, however, that Carnap makes use of (what he calls) "similarity circles", which are unlike equivalence classes in that they may overlap. See Carnap (1967, 129–131).

[11] A relation is considered definable here if we can define it in terms of concepts that we already possess. Note that best approximations to *finite* relations are always definable, provided that we have names for all individuals in the domain.

[12] In the sequel, we will always assume that the R's that we want to approximate are reflexive and symmetrical. Often we will not even bother to mention this.

[13] As will be illustrated in Section 4.3, complications arise when the quantitative approach is applied naively to the case of denumerably infinite graphs. However, in that same subsection we will also show a way of avoiding such complications.

[14] The intransitivity of interbreeding is briefly discussed in Williamson (1990, 114–115).

[15] See Williamson (1990, Chapter 5, Section 5.2).

[16] In the end, Williamson remains neutral on which of these views is to be preferred. See Williamson (1990, Chapter 5, Section 5.2, esp. 79). Nevertheless, Williamson regards the ignorance view for at least some criteria of identity as a "scarcely credible option" (1990, 133). However, in the light of his more recent epistemic theory of vagueness (Williamson 1994) it seems reasonable to expect that Williamson would at present display much more sympathy toward the ignorance view.

[17] Timothy Williamson brought this to our attention.

[18] This example is due to Steven Delvaux.

REFERENCES

Carnap, R.: 1967, *The Logical Structure of the World*, Routledge, London.

Grice, H. P.: 1941, 'Personal Identity', *Mind* **50**, 330–350.

Delvaux, S. and L. Horsten: 2004, 'On Best Transitive Approximations to Simple Graphs', *Acta Informatica* **40**, 637–655.

Krivanek, M. and J. Moravek: 1986, 'NP-Hard Problems in Hierarchical Tree-Clustering', *Acta Informatica* **23**, 311–323.

Williamson, T.: 1986, 'Criteria of Identity and the Axiom of Choice', *Journal of Philosophy* **83**, 380–394.

Williamson, T.: 1990, *Identity and Discrimination*, Blackwell, Oxford.

Williamson, T.: 1994, *Vagueness*, Routledge, London.

Zahn, C. T., Jr.: 1964, 'Approximating Symmetric Relations by Equivalence Relations', *SIAM Journal of Applied Mathematics* **12**, 840–847.

University of Leuven
Center for Logic, Philosophy of Science
& Philosophy of Language
Kardinaal Mercierplein 2
3000 Leuven, Belgium
E-mail: leon.horsten@hiw.kuleuven.ac.de