Notre Dame Journal of Formal Logic
Volume 29, Number 4, Fall 1988

First-Order Logics for Comparative Similarity

TIMOTHY WILLIAMSON

> I am inclined to suspect, from certain data, that the ultimate philosophy of difference and likeness will have to be built upon experiences of intoxication, especially by nitrous oxide gas, which lets us into intuitions the subtlety whereof is denied to the waking state.
>
> —[9], p. 531n

1 If we speak of *degrees* of similarity, what kinds of judgment are we assuming to make sense?

It will be argued that the necessary and sufficient condition for there to be degrees of similarity is that there should be a four-termed relation of comparative similarity—w resembles x at least as much as y resembles z—obeying certain constraints. Of course, nothing turns on how we use the *words* 'degree of similarity'. Rather, the point is to distinguish the different levels of ideological commitment (in Quine's sense) which different kinds of judgment of similarity involve.

In recent years the concept of similarity *tout court* has suffered important vilification (most notably at the hands of Wittgenstein and of Goodman [7]), whose upshot is frequently the claim that it should be replaced by the family of concepts of similarity in various respects. This paper is neutral over such claims; its results are equally valid for judgments of similarity *tout court* and for judgments of similarity in a given respect. What is more, however little our judgments of similarity reflect independently existing structures, we can hardly avoid making them: we should know how much we are letting ourselves in for, and when, in order to understand ourselves (for the psychological ubiquity of similarity, cf. [3], p. 127 and [18], p. 327).

Jean Nicod writes: "A relation which admits degrees is a three-termed relation, for example: '*a* resembles *b* more than *c*'" ([13], p. 61). David Lewis suggests an opposed view: he writes as though talk of degrees of similarity is metaphorical unless they can be measured numerically, and he contrasts such

Received September 8, 1986; revised February 16, 1987

a quantitative concept with the qualitative one which the three-termed relation yields ([11], pp. 50–52). Roughly speaking, it will be argued that while Nicod's criterion is too weak, Lewis's is too strong; but what matters is to place the alternatives which they consider in relation to the alternatives which they neglect. For the contrast between cardinal and ordinal concepts of similarity is by no means the only one which needs to be drawn; the contrast between three-termed and four-termed ordinal relations is equally significant. However, we can first usefully prepare the ground by considering the contrasts which Lewis draws between his cardinal and ordinal concepts.

In Lewis's notation, '$i <_j k$' means that i is more similar than k to j. Suppose that the similarity of i to j could be measured by a real number $d(i,j)$, where the smaller the number, the greater the similarity. Then '$i <_j k$' could be defined as: $d(i,j) < d(k,j)$ (where $<$ is just the usual ordering of the real numbers). Lewis asks what constraints on the three-termed relation are involved in the assumption that there is a real-valued function d in terms of which it can be defined as above, constraints which it would be unnatural to impose on the basis of consideration of the three-termed relation alone. He discusses two candidates.

First, the existence of such a d limits the order-type of comparative similarity. For if we keep j fixed, while allowing i and k to vary, $i <_j k$ is just an order relation on the terms of the similarity relation, and $d(i,j)$ is an order-preserving mapping of them into the real numbers. Thus, given the existence of such a d, there cannot be a set with more members than there are real numbers, no two of whose members resemble j to the same degree (when both $i <_j k$ and $k <_j i$ would fail). Lewis says of this general limitation that it 'hardly seems serious' ([11], p. 51). However, someone might say, for example, that for any infinite cardinals c, c', and c'', if $c < c' < c''$ then c' is more similar than c'' to c; their claim may be false, but it does not seem to be self-contradictory, as a real-valued measure of similarity would make it.[1] Lewis, of course, is concerned specifically with similarity relations amongst possible worlds: yet even here one can argue—as plausibly as the topic permits—that the degrees of similarity between possible worlds, on Lewis's conception of them, are not order-isomorphic to any set of real numbers under their usual ordering. We can restrict attention to worlds whose instants of time can be treated as the real numbers and which are in one of only two states at a given time, off or on (cf. Lewis [12], p. 118). To avoid the complications of coordinate transformations, assume that the world is always off before 0, on at 0, on at 1, and always off after 1. Thus the worlds in which we are interested can be represented as subsets of the open interval (0,1), corresponding to the times in this interval when the world is on. As plausible a way as any of treating the difference between two such sets X and Y is as their symmetric difference D(X,Y): the set of instants which are in X but not Y or vice versa. A plausible constraint on the similarity of these worlds is this: if D(X,Y) is a proper subset of D(Z,Y), then X is more similar than Z to Y. By the Axiom of Choice, there is a well-ordering R of (0,1). For x in (0,1), let X(x) be $\{y\colon y \in (0,1) \;\&\; \mathrm{R}(y,x)\}$. Thus $\mathrm{X}(x) \subset \mathrm{X}(y)$ iff $\mathrm{R}(x,y)$, so $\mathrm{D}(\{\,\},\mathrm{X}(x)) \subset \mathrm{D}(\{\,\},\mathrm{X}(y))$ iff $\mathrm{R}(x,y)$, so the sets $\mathrm{D}(\{\,\},\mathrm{X}(x))$ are well-ordered by proper inclusion. Hence, by assumption, the uncountably many sets X(x) are well-ordered by relative similarity to { }. Now it is well known that no uncountable set of reals is well-ordered by their usual ordering. Hence no real-valued measure under the usual ordering of the reals can give the similarity ordering of all possible worlds, given the above assumptions, and we should not make these assumptions trivially self-contradictory. (For two criticisms based on cardinality considerations of Lewis's view of possible worlds, cf. [5], p. 262 and [6]; Lewis's replies are at [12], pp. 104–108 and 101–104 respectively). Thus we have good reason to avoid the assumption that degree of similarity can always be measured by a real number.

Lewis's second point is that, if we assume the existence of such a function d, we shall be tempted to assume that i is always just as similar to j as j is to i, which in turn entails a constraint on the three-termed relation: if $j <_i k$ and $k <_j i$, then $j <_k i$. Lewis's reasoning can be filled in as follows. The natural demand to make of a real-valued function which is to measure distance in either an ordinary space or a quality one is that it should be a *metric*, i.e., that it should obey the following standard conditions: for any terms i, j, and k of the similarity relation:

(M1) $d(i,j) \geq 0$
(M2) $d(i,j) = 0$ iff $i = j$
(M3) $d(i,j) = d(j,i)$
(M4) $d(i,k) \leq d(i,j) + d(j,k)$.

Of these, (M3) is the immediately relevant one. If the antecedents of Lewis's constraint can be expanded in terms of d, they become:

$$d(j,i) < d(k,i)$$
$$d(k,j) < d(i,j).$$

Its consequent becomes:

$$d(j,k) < d(i,k).$$

Since $<$ is transitive, (M3), applied three times over, is clearly what is required to complete the inference from the antecedents to the consequent. Lewis does not mention that this inference is only the first member of an infinite sequence, each member of which can be validated in the same way in terms of d, and no member of which can be derived from the preceding ones. For, in general, (M3) would allow us to infer from the premises

$$\begin{array}{c} i(0) <_{i(1)} i(2) \\ i(1) <_{i(2)} i(3) \\ \cdots\cdots\cdots\cdots\cdots \\ i(n-2) <_{i(n-1)} i(n) \\ i(n-1) <_{i(n)} i(0) \end{array}$$

to the conclusion:

$$i(1) <_{i(0)} i(n).$$

Such inferences are studied in Section 3. Here, we may examine Lewis's attempt to undermine their validity. His argument is restricted to possible worlds but, if sound, would still show the above not to be universally valid principles of inference about similarity. He suggests that the similarity of i to j may differ from the similarity of j to i because when the former is evaluated, different respects of comparison are weighed according to their importance in j, whereas when the latter is evaluated, different respects of comparison are weighed according to their importance in i:

> Thus it can happen that j is more similar than k to i in the respects of comparison that are important at i; k is more similar than i to j in the respects of comparison that are important at j; yet i is more similar than j to k in the respects of comparison that are important at k. ([11], p. 51)

Thus the inference from $j <_i k$ and $k <_j i$ to $j <_k i$ would fail. Lewis admits that one might continue to measure similarity by a real-valued function d while dropping the assumption (M3), but he questions the point of doing so. However, it is quite misleading to present the relativization of similarity to the point of view of a world as though it were an objection to the symmetry of similarity. For if the relativization is required, we should be able to make general sense of the following four-termed relation, whose fourth term is always a world: x is more similar than y to z when respects of comparison are weighed according to their importance in i. Lewis's three-termed relation is then the result of identifying the third and fourth arguments of the four-termed relation. In effect, j plays a dual role in $i <_j k$. This three-place relation may well be the appropriate one for Lewis to use in giving a semantics for counterfactuals. However, it is not appropriate for a general logic of similarity, since it makes sense only when j (the third and fourth term of the four-termed relation) is a possible world, whereas we shall often wish to speak of the comparative similarity of three terms not one of which is a world. To avoid this problem, we must derive our three-termed similarity relation from the four-termed relation by fixing the fourth world term of the latter (perhaps as the actual world), not by identifying it with the third term. But then the objection to symmetry disappears: for if, when the similarity of x to y and the similarity of y to x are evaluated, respects of comparison are weighed according to their importance in a fixed world i, the same in both cases, then surely x is no more and no less similar to y than y is to x. Thus the above principle of inference can be valid when '$i <_j k$' is interpreted as it needs to be for a general logic of similarity. The relativization of similarity to the point of view of a possible world may or may not be required: in any case, it will not be made explicit from now on.[2]

So far, we have been contrasting metrics with three-termed comparisons. This contrast, however, conflates two kinds of difference. For one of the most obvious features of a metric is that it allows four-termed comparisons, such as $d(w,x) < d(y,z)$, which in turn yield three-termed comparisons when their second and fourth terms (x and z) are identified (note that this four-termed relation is quite different from that considered above, in which the fourth term was a world whose only role was to weigh different respects of comparison). On the other hand, the use of four-termed comparisons—'w is more similar to x than y is to z'—does not seem to require the existence of such a metric; it could be a purely qualitative matter. Thus we must ask, of respects in which the quantitative and the three-termed qualitative approach differ, which are contrasts between the one quantitative approach and the two qualitative approaches and which are contrasts between the two four-termed approaches and the one three-termed approach. In a number of respects, the four-termed qualitative approach will turn out to lack both the implausible commitments of the quantitative approach and the expressive limitations of the three-termed qualitative approach. Section 2 contrasts metrics with four-termed comparative similarity relations; Section 3 contrasts four-termed comparative similarity relations with three-termed comparative similarity relations. It will be argued that the fundamental notion of similarity is that of the four-termed comparative relation.

2 Let us first suppose that we have a metric d that measures similarity within some domain. Let '$T(w,x,y,z)$' mean that w resembles x at least as much as y resembles z ('w resembles x more than y resembles z' can then be formalized by '$T(w,x,y,z)$ & $\sim T(y,z,w,x)$'). Thus $T(w,x,y,z)$ iff $d(w,x) \leq d(y,z)$. What can we deduce from this about T? Since $\leq$ is a connected and transitive relation on the reals, we have automatically:

(T1) $T(w,x,y,z) \vee T(y,z,w,x)$
(T2) $T(u,v,w,x) \rightarrow (T(w,x,y,z) \rightarrow T(u,v,y,z))$.

(As always, universal quantifiers are understood but suppressed.) These are not the only consequences of the fact that d is a real-valued function: we have already seen how it constrains the order-type of degrees of similarity. However, by the Löwenheim-Skolem theorems, such consequences cannot all be stated in first-order formulas whose only atomic predicate (other than identity) is T. Indeed, all the first-order consequences of the fact that d is a real-valued function can be derived from (T1) and (T2). We can show this as follows. Suppose, for a *reductio ad absurdum*, that there is a first-order sentence A whose only atomic predicates are T and identity, which does not follow from (T1) and (T2), and which is true in any model in which T is determined as above by some real-valued function d. Since the set $\{$(T1),(T2),$\sim A\}$ is consistent, it has a countable (finite or infinite) model (X,T) by the Completeness Theorem, where X is a set and T a four-termed relation on X. Define a binary relation R on X^2, the Cartesian product of X with itself, by: $R((w,x),(y,z))$ iff $T(w,x,y,z)$ and $T(y,z,w,x)$. Then T is an equivalence relation, for (T1) makes it reflexive, its form makes it symmetric, and (T2) makes it transitive. Let $|x,y|$ be the R-equivalence class of (x,y). It is easy to check that a relation $\leq$ on the R-equivalence classes can be consistently defined by: $|w,x| \leq |y,z|$ iff $T(w,x,y,z)$. $\leq$ is a connected, antisymmetric, transitive relation: that is, a reflexive total order (cf. [10], pp. 14–16). Its field is countable, since X^2 is countable. As is well known, any totally ordered countable set is order-isomorphic to a subset of the rationals and therefore of the reals. Let o be such an order isomorphism, and put $d(x,y) = o(|x,y|)$. Then $d(w,x) \leq d(y,z)$ iff $T(w,x,y,z)$. Thus A should be true in (X,T), which is a contradiction. Hence (T1) and (T2) do exhaust the first-order consequences of the definition of T in terms of a real-valued function d. We must now examine the consequences of d's being a metric.

(M1) and (M2) together say that x and y are maximally similar when, and only when, they are identical. That is, when T is defined in terms of d, they entail:

(T3) $T(x,x,y,z)$
(T4) $T(x,y,y,y) \rightarrow x = y$;

and (M3) obviously entails:

(T5) $T(x,y,y,x)$.

(T3)–(T5) exhaust the extra consequences (of any order) of d's being a metric, in the following sense: if there is a real-valued function d (not necessarily a metric) such that $T(w,x,y,z)$ iff $d(w,x) \leq d(y,z)$, for all w, x, y and z (in the domain), and T obeys (T3)–(T5), then there is a metric d' such that $T(w,x,y,z)$ iff $d'(w,x) \leq d'(y,z)$, for all w, x, y and z. We can prove this as follows. Suppose that there is such a real-valued function d, and that T obeys (T3)–(T5). Thus by (T3) and (T4) respectively:

(M5) $d(x,x) \leq d(y,z)$
(M6) $d(x,y) \leq d(y,y) \rightarrow x = y$.

(T5) obviously entails that d obeys (M3). Thus it suffices to show that if a real-valued function d obeys (M3), (M5), and (M6), then there is a metric d' such that $d(w,x) \leq d(y,z)$ iff $d'(w,x) \leq d'(y,z)$, for all w, x, y and z. First, we satisfy (M1) and (M2); for any v, put $d''(x,y) = d(x,y) - d(v,v)$. Clearly, d'' obeys (M1)–(M3) and $d(w,x) \leq d(y,z)$ iff $d''(w,x) \leq d''(y,z)$. Thus it suffices to show that if a real-valued function d'' obeys (M1)–(M3), then there is a metric d' such that $d''(w,x) \leq d''(y,z)$ iff $d'(w,x) \leq d'(y,z)$. Such a d' can be defined as follows:

If $d''(x,y) = 0$, $d'(x,y) = 0$
If $d''(x,y) \neq 0$, $d'(x,y) = 2 - 1/(1 + d''(x,y))$.

One can easily check that d' obeys (M1)–(M3). It obeys (M4) because if $d'(x,y) \neq 0$, $1 < d'(x,y) < 2$; now if $d'(x,y) = 0$ or $d'(y,z) = 0$ or $d'(x,z) = 0$, $d'(x,z) \leq d'(x,y) + d'(y,z)$ is trivial, while otherwise $d'(x,z) < 2 = 1 + 1 < d'(x,y) + d'(y,z)$. Moreover, $d''(w,x) \leq d''(y,z)$ iff $d'(w,x) \leq d'(y,z)$. QED. Putting this result together with that of the previous paragraph: (T1)–(T5) exhaust the first-order consequences of the assumption that T is definable in terms of a metric.

A point worth noting emerged in the last paragraph: that (M4)—the "triangle inequality"—entails no extra constraint on T whatsoever (cf. [3], p. 130). That is, if some real-valued function d'' which obeys (M1)–(M3) is such that $T(w,x,y,z)$ and $d''(w,x) \leq d''(y,z)$ are equivalent, then some real-valued function d' which obeys (M1)–(M4) is such that $T(w,x,y,z)$ and $d'(w,x) \leq d'(y,z)$ are equivalent. Now (M4) is the only one of the conditions on a metric which uses the additive (or multiplicative) structure of the real numbers, and that structure is the source of many of the doubts about the numerical measurement of similarity. For one might very well regard such judgments as 'w resembles x as much as y resembles z' as meaningful while taking the opposite attitude towards such judgments as 'w resembles x twice as much as y resembles z' and 'the similarity of w to x exceeds the similarity of y to z by more than the similarity of w_0 to x_0 exceeds the similarity y_0 to z_0'. Since this extra structure, which a metric makes available, has no implications for the relation T, we automatically neutralize doubts about it by switching attention from d' to T (cf. [18], p. 329).

Essentially the same point can be made by the observation that if T does correspond to a metric, it corresponds to many, which (with trivial exceptions) are not even constant multiples of each other. For suppose that T does correspond to a metric d'' (that is, $T(w,x,y,z)$ when and only when $d''(w,x) \leq d''(y,z)$); then if d' is defined in terms of d'' as before, d' will also be a metric to which T corresponds ($T(w,x,y,z)$ when and only when $d'(w,x) \leq d'(y,z)$); yet d' is unlikely to be a linear function of d''. The mapping of quantitative metrics to four-termed qualitative relations is many-one; it cannot be inverted. However, if T avoids the spurious precision of d, it must be admitted that the argument cuts both ways. For there is something right about the idea, which the triangle inequality expresses, that if we know upper limits to the difference between x and y and the difference between y and z then we can infer an upper limit to the difference between x and z; this insight cannot be captured in terms of the relation T. However, it will not be pursued here.

Each of (T1)–(T5) can be questioned as a condition on comparative similarity.

(T1): If w and x are things of one kind and y and z of another, why should the similarity of w to x not be incommensurable with the similarity of y to z? One form of skepticism about similarity could take the form of holding (T1) to fail in many or most cases; it would have the advantage of making similarity a more or less useless relation in philosophy, while not being committed to the implausible claim that judgments of similarity are literally meaningless.

(T2): We can adapt Condorcet's paradox from the theory of social choice ([2], p. 122) to construct a putative counterexample to (T2). Suppose that we are to combine three apparently incommensurable respects of similarity into a concept of overall similarity on some domain; comparative similarity in each of these respects is assumed to be transitive. We decide that w overall-resembles x (at least) as much as y overall-resembles z iff there are at least as many of the three respects in which w resembles x more than y resembles z as there are in which y resembles z more than w resembles x. Now suppose that in respect 1, u resembles v more than w resembles x and w resembles x more than y resembles z; in respect 2, w resembles x more than y resembles z and y resembles z more than u resembles v; in respect 3, y resembles z more than u resembles v and u resembles v more than w resembles x. Then, by our criterion, u overall-resembles v as much as w overall-resembles x (respects 1 and 3 vs. respect 2) and w overall-resembles x as much as y overall-resembles z (respects 1 and 2 vs. respect 3), but u does not overall-resemble v as much as y overall-resembles z (respect 1 vs. respects 2 and 3).

(T3): Tversky argues ([18], p. 328) that if the similarity of stimuli x and y is measured by the probability that a subject, when presented with x and then with y, will judge the stimuli to have been identical, (T3) may sometimes fail.

(T4): (T4) amounts to the Identity of Indiscernibles, for if x and y are qualitatively identical, x is as similar to y as y is to itself, and (T4) then says that they are numerically identical. The corresponding condition on the metric d is that $d(x,y) = 0$ entails $x = y$. If one drops this half of (M2), its less controversial consequences can still be deduced from (M4). That is, (M4) entails:

(M7) $d(x,y) = 0 \rightarrow d(x,z) = d(y,z)$.

Expressed in terms of T, this becomes:

(T6) $T(x,y,y,y) \rightarrow T(x,z,y,z)$.

That is, if x and y are perfectly similar to each other, then each is as similar as the other to anything else, a much less controversial claim.

(T5): We have already discussed doubts which can be raised about (T5).

In spite of all this, we shall continue to work in terms of (T1)–(T5). We can be fairly sure that they generate the strongest plausible first-order theory of the *logic* of comparative similarity, for any first-order principle A about similarity which is supposed to be logically valid should at least be valid in all those domains in which similarity can be measured numerically; but no metric seems to be excluded from measuring similarity on purely formal grounds, and thus such a principle would have to hold in all models in which T corresponds to some metric; but then, as we saw above, A would have to follow from (T1)–(T5). Thus (T1)–(T5) entail any plausible first-order theory of the *logic* of comparative similarity, if not vice versa. Now most of the results which are proved below about (T1)–(T5) can easily be extended to weakenings of them, so that concentration on (T1)–(T5) is not a major limitation; we are not committed to their plausibility as principles about similarity. (The purely universal axioms of [3], at p. 130, are equivalent to (T1)–(T5); they are also in effect clauses 1 and 2 of the definition of an *absolute-difference structure* at [10], p. 172, the other clauses of which depend on the assumption of a unidimensional domain.)

Moreover, the doubts about (T2) and (T3) seem to stem from an overly operationalistic approach to similarity, and will not be further pursued, while systems without (T5) are not rich enough to raise the issues with which we shall mostly be concerned. Thus the two weakenings of (T1)–(T5) on which we concentrate below are the dropping of (T1) and the replacement of (T4) by (T6).

It was claimed above that the relation T allows us to speak of *degrees* of similarity. We have already seen the main idea. A binary relation R is defined on ordered pairs by: $R((w,x),(y,z))$ iff $T(w,x,y,z)$ and $T(y,z,w,x)$. R is symmetric by its form. The relatively uncontroversial (T2) makes it transitive. Its reflexivity is equivalent to:

(T7) $T(x,y,x,y)$.

(T7) is unproblematic, and in any case follows from (T2) and (T5). Thus we can unproblematically assume R to be an equivalence relation, and thus (if we wish) identify degrees of similarity with R-equivalence classes. As before, $|x,y|$ is the R-equivalence class of (x,y) and we can put $|w,x| \leq |y,z|$ iff $T(w,x,y,z)$. $|w,x| < |y,z|$ iff $|w,x| \leq |y,z|$ without $|y,z| \leq |w,x|$. $\leq$ is then a reflexive, antisymmetric, transitive relation, i.e., a reflexive partial order. (T2) would make it connected, i.e., a reflexive total order. By (T3), (x,x) and (y,y) are always

R-related; let their equivalence class be, with suggestive ambiguity, 0. Then (T3)–(T5) can be restated as:

(T3′) $0 \leq |y,z|$
(T4′) $|x,y| = 0 \rightarrow x = y$
(T5′) $|x,y| = |y,x|$.

(Compare conditions (M1)–(M3) on a metric). Similarly, (T6) becomes:

(T6′) $|x,y| = 0 \rightarrow |x,z| = |y,z|$.

Thus at least some of the structure of a metric can be recovered. The reader is left to consider whether further plausible constraints, in higher-order terms, could be imposed on the system of degrees of similarity.

3 We may now turn to the contrast between four-termed and three-termed comparative similarity relations. First it will be argued that the contrast is of some philosophical interest; then its logical basis will be explored.

In practice, we make comparisons far more often amongst three terms than we do amongst four; many authors use a three-termed relation without discussion (e.g. [13], [15], pp. 117–135, [4] at p. 4). Sometimes, the three-termed relation may also be empirically more accessible ([10], pp. 177–178, [16], p. 18). Presumably, we do not usually *need* to make four-termed comparisons: although we might explain why John's handwriting is harder to read than Peter's by the fact that the letters of the alphabet as written by John resemble each other more than the corresponding letters do as written by Peter. Lewis gives 'A red thing could resemble an orange thing more closely than a red thing could resemble a blue thing' as an example of an English sentence which is most naturally analyzed by means of a four-termed cross-world similarity relation ([12], p. 13). More generally, any concept of *homogeneity* which permits the homogeneity of a set to be compared with the homogeneity of a disjoint set in some respect will presuppose the possibility of four-term comparisons in that respect (between two members of one set and two members of the other).

The difference between three-termed and four-termed similarity relations can also be of philosophical significance. Here is a rather speculative example. Hilpinen has in effect shown how, given a three-termed relation of comparative similarity on possible worlds such as David Lewis employs in his semantics for counterfactuals, one can define concepts of the *closeness to the truth* and *informativeness about the truth* of a proposition ([8], pp. 27–32 and [12], pp. 24–27; Hilpinen uses Lewis's method of spheres, which is equivalent to using the three-termed relation). Such concepts are obviously of potential importance for a realist philosophy of science. For any consistent propositions P and Q, Hilpinen's definitions amount to the following:

> P is at least as close to the truth as Q in a world i iff, for every world k at which Q is true, there is a world j at which P is true such that j is at least as similar to i as k is.
>
> P is at least as informative about the truth as Q in a world i iff, for every world j at which P is true, there is a world k at which Q is true such that j is at least as similar to i as k is.

Now suppose that one is investigating the not implausible hypothesis that how much can be known about the laws of nature depends on *which* laws of nature obtain (some laws make science easy, some make it hard). Let j and k be worlds in which different laws of nature obtain, and let P and Q be well-confirmed theories in j and k, respectively. One will be interested in comparing the merits of P as a theory about j with those of Q as a theory about k. Such merits would include closeness to the truth and informativeness about the truth. However, Hilpinen's definitions give no sense to judgments of the forms 'P is at least as close to the truth in h as Q is to the truth in i' and 'P is at least as informative about the truth in h as Q is about the truth in i'. If only a three-termed similarity relation is permitted, there is no natural way of extending his definitions to give sense to such judgments. In contrast, suppose that one is permitted a four-termed similarity relation; then one can extend Hilpinen's definitions as follows:

> P is at least as close to the truth in a world h as Q is to the truth in a world i iff, for every world k at which Q is true, there is a world j at which P is true such that j is at least as similar to h as k is to i.
>
> P is at least as informative about the truth in a world h as Q is about the truth in a world i iff, for every world j at which P is true, there is a world k at which Q is true such that j is at least as similar to h as k is to i.

Thus philosophical questions can turn on the difference between three-termed and four-termed similarity relations. The eventual upshot of the discussion below will be that three-termed similarity relations are conceptually derivative from corresponding four-termed relations. Hence, if Hilpinen's original definitions make sense, so do the above extensions of them, which would considerably enhance the analytic power and usefulness of the concepts of closeness to, and informativeness about, the truth. Psychologically, it is easy to understand why we should find four-termed comparisons more confusing than we do three-termed ones ([9], p. 531n); the logical relations between them are what concern us here.

We have seen that a four-termed relation yields at least a rudimentary theory of degrees of similarity. Conversely, if we can speak of degrees of similarity, we can at once construct four-termed relations by comparing the degree of similarity of w to x with the degree of similarity of y to z, for talk of degrees of similarity would make no sense unless they were subject to some order relation. Thus a theory of similarity based on a three-termed relation allows the construction of degrees of similarity iff a suitable four-termed relation can be defined in terms of the three-termed relation. (For another construction of four-termed from three-termed comparison, requiring much stronger assumptions, cf. [10], pp. 178–195.)

Given the four-termed comparative similarity relation T, we define a three-termed comparative similarity relation S in terms of it by:

(Def) $S(x,y,z) \leftrightarrow T(x,y,z,y)$.

In other words, x is at least as similar as z to y iff x is at least as similar to y as z is to y. Note that, in the absence of (T2) or (T5), (Def) differs nontrivially from:

(Def′) $S(x,y,z) \leftrightarrow T(x,y,y,z)$.

For, given (Def), $S(x,y,x)$ expresses the anodyne (T7) ($T(x,y,x,y)$), whereas, given (Def′), it expresses the controversial (T5) ($T(x,y,y,x)$) itself. A thorough investigation of the logic of nonsymmetric similarity would treat the distinct three-termed relations introduced by (Def) and (Def′) ('x is more similar than z to y' and 'x is more similar to y than y is to z') separately. Since we are largely concerned with systems that include (T2) and (T5), we use only (Def) in the following.

We can easily show that several four-termed relations, all obeying the appropriate constraints, can give rise to the same three-termed relation, so that the definition of the three-termed relation in terms of a four-termed relation cannot in general be inverted. If we do define a four-termed relation in terms of a three-termed one, we shall have chosen the former from the alternatives in a more or less arbitrary way.

Consider an example. Suppose that we have a domain of four distinct objects, a, b, c, and d, whose degrees of similarity (as defined in terms of T) are, in order of decreasing similarity:

$$\begin{aligned}
|a,a| &= |b,b| = |c,c| = |d,d| \\
|a,b| &= |b,a| \\
|c,d| &= |d,c| \\
|a,c| &= |c,a| = |b,d| = |d,b| \\
|a,d| &= |d,a| = |b,c| = |c,b|.
\end{aligned}$$

It is clear that the three-termed relation to which this ordering gives rise could also have been obtained from the ordering:

$$\begin{aligned}
|a,a| &= |b,b| = |c,c| = |d,d| \\
|c,d| &= |d,c| \\
|a,b| &= |b,a| \\
|a,c| &= |c,a| = |b,d| = |d,b| \\
|a,d| &= |d,a| = |b,c| = |c,b|,
\end{aligned}$$

for the two orderings differ only with respect to the relative similarities of a to b and c to d. Since this is a comparison between two pairs (neither of the form (x,x)) without a common member, it cannot be represented by the three-termed relation. The two orderings correspond to incompatible claims about T: the first would assert $T(a,b,c,d)$, the second would deny it. Thus there is no road back from the three-termed relation to the four-termed relation. In case it is doubted that these orderings are possible for degrees of *similarity*, we could imagine that the objects are being compared with respect to two quantifiable dimensions, so that each object x can be represented as a pair of real numbers $(f(x),g(x))$. Assume that similarity can be measured in the standard Euclidean way, in other words that $|x,y|$ corresponds to the real number $\sqrt{(f(x)-f(y))^2+(g(x)-g(y))^2}$. Then if we put $(f(a),g(a)) = (0,0), (f(b),g(b)) = (3,0), (f(c),g(c)) = (1,5)$ and $(f(d),g(d)) = (2,5)$, we obtain the second ordering, while if we interchange a with c and b with d in these assignments, we obtain the first. Thus they can hardly be objected to as similarity orderings on formal grounds.

We now come to the comparison of the possible conditions on T, (T1)–(T7), with possible conditions on the three-termed relation S. In one direction, the problem can be simply posed. Given a set of sentences in which 'T' but not 'S' appears, what sentences do they, together with (Def), entail in which 'S' but not 'T' appears? However, if we merely reverse 'S' and 'T' in this question, the result is not a very interesting approach to the opposite direction. For it is consistent with (Def) and any consistent set of sentences in which 'S' but not 'T' appears that '$T(w,x,y,z)$' should be true only when $x = z$; thus even the innocuous (T3) ($T(x,x,y,z)$) could not be deduced in this way; nor could (T1) or (T5). A looser notion of correspondence between conditions on S and conditions on T would be more appropriate.

Such a notion may be defined as follows. Let $L(S)$ be the first-order language whose atomic predicates are '=' and (the three-termed) 'S', $L(T)$ the first-order language whose atomic predicates are '=' and (the four-termed) 'T', and $L(S, T)$ the first-order language whose atomic predicates are '=', 'S', and 'T'. '=' is always interpreted as identity (so we shall not mention it when we specify a model); the appropriate axioms for it will be taken as part of the underlying logic. Let A be a set of sentences in $L(S)$ and B a set of sentences in $L(T)$. We say that A *corresponds to* B iff, for every $L(S)$-model (X,S) of A, there is a relation T on X such that (X,S,T) is an $L(S,T)$-model of $B \cup \{(\mathrm{Def})\}$. Similarly, B *corresponds to* A iff, for every $L(T)$-model (X,T) of B, there is a relation S on X such that (X,S,T) is an $L(S,T)$-model of $A \cup \{(\mathrm{Def})\}$. A and B *correspond simpliciter* iff A corresponds to B and B corresponds to A; in other words, any three-termed relation which obeys A is the restriction of a four-termed relation which obeys B, and any four-termed relation which obeys B can be restricted to a three-termed relation which obeys A. We shall see that this notion articulates the intuitive idea that a theory about S may correspond to a theory about T even when the former, together with (Def), does not entail the latter. Examples will be given below.

Where A and B are as above, it is easy to see that B corresponds to A iff $B \cup \{(\mathrm{Def})\}$ entails (each sentence in) A (that is, every model of $B \cup \{(\mathrm{Def})\}$ is a model of A). For since (Def) defines S in terms of T, there is a function f such that (X,S,T) is a model of (Def) iff S = f(T). Now for every model (X,f(T),T) of $B \cup \{(\mathrm{Def})\}$, (X,T) is an $L(T)$-model of B, so if B corresponds to A, there is a model (X,S,T) of $A \cup \{(\mathrm{Def})\}$, but then f(T) = S, so (X,f(T),T) is a model of A; thus $B \cup \{(\mathrm{Def})\}$ entails A. Conversely, if $B \cup \{(\mathrm{Def})\}$ entails A, for every $L(T)$-model (X,T) of B, (X,f(T),T) is a model of $B \cup \{(\mathrm{Def})\}$ and so of $A \cup \{(\mathrm{Def})\}$; thus B corresponds to A.

It follows that, given a theory B in $L(T)$, there is at most one theory A in $L(S)$ such that A and B correspond (a theory here is a deductively closed set of first-order sentences). For suppose that A and A' are theories in $L(S)$ such that A and B correspond and A' and B correspond. By the previous paragraph, $B \cup \{(\mathrm{Def})\}$ entails $A \cup A'$. Let (X,S) be an $L(S)$-model of A; since A corresponds to B, there is a relation T on X such that (X,S,T) is an $L(S,T)$-model of $B \cup \{(\mathrm{Def})\}$; since $B \cup \{(\mathrm{Def})\}$ entails A', (X,S,T) is an $L(S,T)$-model of A', so (X,S) is an $L(S)$-model of A'. Thus A entails A'; similarly, A' entails A; since A and A' are theories, $A = A'$.

We now need some examples of correspondence. A theory in $L(S)$ will be presented such that it and the maximal logic of the four-termed similarity relation, generated by (T1)–(T5), correspond. The correspondence of various pairs of weaker theories can be proved by minor modifications of the proof for this case. Consider the following axioms:

(S1) $S(x,y,z) \vee S(z,y,x)$

(S2) $S(x,y,y) \rightarrow x = y$

(S3$_n$) $(S(v_0,v_1,v_2) \& \ldots \& S(v_{n-2},v_{n-1},v_n) \&$
$S(v_{n-1},v_n,v_0)) \rightarrow S(v_1,v_0,v_n)$.

Let (S3) be the infinite schema with each (S3$_n$) as an instance for $2 \leq n$. From (S1) and (S3$_5$) we can derive a more straightforward transitivity principle which will be useful in what follows:

(S4) $S(x,w,y) \rightarrow (S(y,w,z) \rightarrow S(x,w,z))$.

For we can collect an antecedent of the appropriate form for (S3$_5$); thus:

$S(w,x,w)$	(from (S1))
$S(x,w,y)$	(assumption)
$S(w,y,w)$	(from (S1))
$S(y,w,z)$	(assumption)
$S(w,z,w)$	(from (S1)).

$S(x,w,z)$ is then the corresponding consequent.[3]

Let (S1)–(S3) and (T1)–(T5) generate the theories (Sa) and (Ta) respectively. Then it can be proved (Theorem I in the Appendix) that (Sa) and (Ta) correspond. It follows, by what was noted above, that (Sa) is the only theory in $L(S)$ such that it and (Ta) correspond. However, (Ta) is not the only theory in $L(T)$ such that (Sa) and it correspond. For example, let (Ta′) be the theory generated by (T1), (T2), (T4), and (T5); then (Sa) and (Ta′) correspond. For (T3) is not used in the proof that (Ta) corresponds to (Sa), so it proves equally that (Ta′) corresponds to (Sa); conversely, since (Sa) corresponds to (Ta) and every model of (Ta) is a model of (Ta′), (Sa) corresponds to (Ta′) (it is proved in the Appendix that (T3) is independent of the other axioms of (Ta)). More trivially, if we merely replace each component of the form '$S(v,v',v'')$' in the axioms of (Sa) by one of the form '$T(v,v',v'',v')$', (Sa) and the result obviously correspond, even though the result does not require that '$T(w,x,y,z)$' ever holds when $x \neq z$, and thus does not entail (T1), (T3), or (T5). This construction can be applied to any theory A in $L(S)$: call the result $T(A)$ (we could even add '$T(w,x,y,z) \rightarrow x = z$' as an axiom). However, this is only to be expected, given that three-termed similarity does not uniquely determine four-termed similarity; moreover, amongst the serious candidates for a logic of four-termed similarity which we shall examine, (Ta) is the only one which stands in this relation to (Sa).

The result that (Sa) corresponds to (Ta) can also be thought of as a *representation theorem*, saying that any three-termed relation which obeys (Sa) can be derived by (Def) from a four-termed relation which obeys (Ta), so that $S(x,y,z)$ can be represented in terms of degrees as $|x,y| \leq |z,y|$. It is then natural to ask when this representation is *unique*, that is, when a given relation obeying (Sa) can be so derived from only one relation obeying (Ta), up to extensional equivalence. To state the answer, a relation Q is defined on ordered pairs, which relates (x,x) to (y,z), (x,y) to itself, (x,y) to (y,z) when $S(x,y,z)$, and does not relate any other pairs. Then the representation is unique iff the ancestral of Q is connected (Theorem I′).

We noted above that (Ta) may be too strong to be a plausible logic of similarity. The obvious weakenings of it, we noted, were to drop (T1) and to replace (T4) by (T6). Applying one or both of these, we have three weaker theories: (Tb) generated by (T2)–(T5), (Tc) generated by (T1)–(T3) and (T5)–(T6), and (Td) generated by (T2), (T3), (T5), and (T6). In order to find correspondences between these theories and theories in $L(S)$, we state three more conditions on S:

(S5) $S(x,y,x)$
(S6) $S(x,x,y)$
(S7) $S(x,y,y) \rightarrow S(y,z,x)$.

It is an easy exercise to check that the following deductive relations hold:

(S3),(S5) $\vdash$ (S4)
(S1) $\vdash$ (S5)
(S6),(S7) $\vdash$ (S5)
(S1),(S7) $\vdash$ (S6)
(S2),(S5) $\vdash$ (S7).

Now let (S2), (S3), (S5), and (S6) generate the theory (Sb), let (S1), (S3), and (S7) generate (Sc) and let (S3), (S6), and (S7) generate (Sd). Using the above relations, one can check that each of (Sa), (Sb), (Sc), and (Sd) entails each of (S4), (S5), (S6), and (S7) (that the above axiomatizations of (Sa)–(Sd) and (Ta)–(Td) are all independent is proved in the Appendix). Similarly, each of (Ta), (Tb), (Tc), and (Td) entails (T7), since it follows from (T2) and (T5). Then it can be proved that (Sb) and (Tb) correspond (Theorem II), (Sc) and (Tc) correspond (Theorem III), and (Sd) and (Td) correspond (Theorem IV).

It can also be proved (Theorems II′, III′, and IV′) that the condition for the representability of a three-termed relation as derived from a four-termed one, both obeying the appropriate constraints, remains the same as before—the connectivity of the ancestral of the relation Q. The proofs of Theorems I′–IV′ also show that if (X,S) is a model of (Sb) ((Sd)) but not of (S1), it has more than one extension to a model of (Tb) ∪ {(Def)} ((Td) ∪ {(Def)}). In other words, if one is seeking to determine degrees of similarity uniquely from three-termed comparative similarity, there is nothing to be *lost* by postulating (S1), that comparative similarity is connected.

In the same way as before, we note from these correspondences that, for any theory A in $L(S)$: if A and (Tb) correspond then A is (Sb); if A and (Tc) correspond then A is (Sc); if A and (Td) correspond then A is (Sd). On the other hand, we can as before trivially construct theories B, B', and B'' in $L(T)$ such that: (Sb) and B correspond but B is not (Tb); (Sc) and B' correspond but B' is not (Tc); (Sd) and B'' correspond but B'' is not (Td). For we can always put $B = T$((Sb)), $B' = T$((*Sc*)), $B'' = T$((Sd)); then none of these theories will entail (T5). Since this construction is always available, it is no surprise when, given a theory A in $L(S)$, we can find a theory B in $L(T)$ such that A and B correspond. However, we were working the other way round: we formulated some theories in $L(T)$, not with an eye to $L(S)$ but because of their relation to the conditions on a metric, and then discovered that there are corresponding theories in $L(S)$. Was that a lucky accident, or is it the general rule? In other words, given a theory B in $L(T)$, is there always a theory A in $L(S)$ such that A and B correspond? By answering this question we can deepen our understanding of the difference between using a four-termed similarity relation and using a three-termed one.

We may first note that, given a theory B in $L(T)$, there is only one candidate for a theory A in $L(S)$ such that A and B correspond. For let $S(B)$ be the theory which consists of all consequences of $B \cup \{(\text{Def})\}$ in $L(S)$: then if A is a theory in $L(S)$ such that A and B correspond, A in $S(B)$. For if B corresponds to A, $B \cup \{(\text{Def})\}$ entails A, so $A \subseteq S(B)$, so every model of $S(B)$ is a model of A; conversely, if A corresponds to B, for every model (X,S) of A there is a model (X,S,T) of $B \cup \{(\text{Def})\}$ and therefore of $S(B)$, so that (X,S) is also a model of $S(B)$. Consequently, by Theorems I–IV, S((Ta)) = (Sa), S((Tb)) = (Sb), S((Tc)) = (Sc) and S((Td)) = (Sd). Our question has now become: given a theory B in $L(T)$, must $S(B)$ and B correspond?

We can give an affirmative answer to the question only if we impose some restriction on B. For consider the sentence:

(T8) $\exists w \exists x \exists y \forall z \sim T(x,y,w,z)$.

It says, roughly, that there are three things of which the first does not differ from anything by as much as the other two differ from each other. Let (Te) be the theory generated by (T1)–(T5) and (T8). Then it can be proved (Theorem VI) that there is no theory A in $L(S)$ such that A and (Te) correspond. However, note that (T8) makes an existential claim. In contrast, the axioms of all the other theories in $L(T)$ (and in $L(S)$) which we have considered are all purely universal in form: each consists of an unquantified matrix to which universal quantifiers are implicitly prefixed. This is as we should expect, for the variables range not over degrees of resemblance but simply objects of any kind, the resemblances amongst which we wish to speak of. Thus it would be inappropriate for a logic of resemblance to make any existence claims about such objects.[4] Given that the theory B obeys this condition—that it should be axiomatizable by purely universal sentences—it can be proved (Theorem V) that $S(B)$ and B do correspond.

The fact that no set of sentences, finite or infinite, in the first-order language of the three-termed relation can express the theory (Te), even in the weak sense of correspondence (let alone translation), is further evidence that the three-termed relation should not be taken as basic. For the critical sentence (T8) seems to be as intelligible as any other claim of comparative similarity.

To summarize the results of this section: The difference between three- and four-termed similarity relations has significant ramifications, for instance, for the analysis of verisimilitude. The four-termed relation can be recovered from the three-termed one only in special circumstances. Interesting theories about the former are not equivalent to theories about the latter. However, if a theory of the four-termed relation is axiomatizable by purely universal sentences, any model of its consequences for the three-termed relation can be expanded to a model of the original theory. Even this result can fail when the condition about universal sentences is dropped. Since four-termed comparisons can be made by means of the nonmetric concept of degrees of similarity, the latter is considerably more powerful than the concept of the three-termed relation of comparative similarity.

4 We now turn to investigate a striking difference among the theories we have investigated, between those in $L(T)$ and those in $L(S)$, for all the axiomatizations of theories about the four-termed relation which we have looked at involved only finitely many axioms, whereas all the axiomatizations of theories about the three-termed relation which we have looked at involved infinitely many axioms, since they all used the axiom schema (S3). Indeed, it can be proved (Theorem VII) that the infinitely many instances of (S3) are not equivalent to a finite number of axioms.

Plausible theories of the four-termed comparative similarity relation are finitely axiomatizable; the corresponding theories of the three-termed comparative similarity relation are not. This fact can be used to support a conclusion already suggested by grammatical indicators ('x is more similar than y to z' looks like a transformation of 'x is more similar to z than y is similar to z', which itself seems to result by identification of arguments from 'x is more similar to z than y is similar to w'): that the four-termed relation is conceptually more basic than its three-termed counterpart. It is not that there is anything *wrong* with a theory just because it is not finitely axiomatizable; any interesting logic has either an axiom schema or (what is no *more* unproblematic) an inference schema with in either case infinitely many instances (cf. [19]). The point is instead a more specific one. (S3) is a *complicated* schema; it is not immediately and self-evidently valid (on its intended interpretation), yet if we ask ourselves whether it is valid, or plausible, sufficient thought will convince us that it is (at least, it convinces the present author). Thus the validity, or apparent validity, of the schema cries out for some kind of explanation; moreover, since our knowledge of it seems to be inferred rather than immediate, the best way of explaining it would be to deduce the instances of the schema from assumptions which we seem to accept without apparently making such complicated inferences. There is little prospect of doing this in terms of assumptions about the three-termed relation. For the strongest remotely plausible assumption that we can make about the three-termed relation *in general* (and we are trying to explain the general validity of (S3)) is that it is derived from some metric (as $d(x,y) \leq d(z,y)$): but then it will be related by (Def) to a four-termed relation ($d(w,x) \leq d(y,z)$), where (Ta) exhausts what we can assume about the latter, so that (Sa) exhausts what we can assume in logic about the former (by Theorem I). Since (Sa) cannot be finitely axiomatized, there is little chance of using its theorems to derive all instances of (S3) from a satisfying simpler basis. As was noted in the Corollary to Theorem VII, the same goes for theories about the three-termed relation which are weaker than (Sa). On the other hand (T2) and (T5)—together with the "bridge law" (Def)—are two plausible assumptions about the four-termed relation, as simple as such assumptions can be, from which we can derive each instance of the schema. By inference to the best explanation, we should therefore assume that our knowledge of the three-termed relation is based on our prior knowledge of the four-termed relation. By Theorem VI, this commits us to the meaningfulness of judgments such as (T8), which cannot be expressed in terms of the three-termed relation; we should embrace this independently plausible consequence. In Section 2 we saw that the four-termed relation allows us to speak of *degrees* of similarity: thus we can conclude that we may speak of degrees of similarity if we may speak of similarity at all.

We arrive at a related point if we evaluate the claim (suggested by the discussion of Lewis) that (S3) expresses the symmetry of similarity. In favor of the claim, we may note that (S3) can be deduced from (Def) and (Ta), but that when (T5), which expresses symmetry straightforwardly, is removed from (Ta), no instance of (S3) can any longer be deduced. For if (T1)–(T4) generate (Tf), the corresponding theory (Sf) in $L(S)$ is generated simply by (S1), (S2), and (S4) (Theorem VIII; the proof of Theorem VII shows that no instance of (S3) follows from (Sf)).

However, we have already noted that (T5) cannot be deduced from (Def) and (S1)–(S3). We can now strengthen this to the following: (Def), (S1)–(S3), and (T1)–(T4) do not entail (T5). For if they did, there would be some k such that (Def), (S1), (S2), $(S3_2), \ldots, (S3_k)$, and (T1)–(T4) entailed (T5), by Compactness. But if, in the proof of Theorem VII, we put $\mathrm{T}(w,x,y,z)$ equivalent to $d(w,x) \leq d(y,z)$, we have a model for (Def), (S1), (S2), $(S3_2), \ldots, (S3_k)$, and (T1)–(T4) in which (T5) is false.

One might try using (Def′) rather than (Def) to link (S3) and (T5) (of course, (Def′), (T2), and (T5) entail (S3)). Now (Def′), (S1)–(S3), and (T1)–(T4) do trivially entail (T5), for (S1) entails (S5), which is equivalent to (T5) by (Def′). However, this clearly fails to show any connection between (S3) and (T5). What we can show is that (Def′), (S2), (S3), (S6), and (T1)–(T4) do *not* entail (T5). For define d on a domain $\{a,b\}$ by: $d(a,a) = d(b,b) = 0$; $d(a,b) = 1$; $d(b,a) = 2$; let $\mathrm{T}(w,x,y,z)$ be equivalent to $d(w,x) \leq d(y,z)$ and define S in terms of T by (Def′). It is easy to see that (S2), (S6), and (T1)–(T4) are true in this model. To see that it makes $(S3_n)$ true, suppose that $\mathrm{S}(v_0,v_1,v_2), \ldots, \mathrm{S}(v_{n-1},v_n,v_0)$ without $\mathrm{S}(v_1,v_0,v_n)$. If $v_0 = v_n$ then, by (S2) and induction, $v_0 = v_1$, whence (S6) would give $\mathrm{S}(v_1,v_0,v_n)$. Thus $\mathrm{S}(v_1,v_0,v_n)$ can fail only if $v_1 = b$, $v_0 = a$, and $v_n = b$. Then $\mathrm{S}(v_0,v_1,v_2)$ requires $v_2 = a$: but this contradicts $\mathrm{S}(v_1,v_2,v_i)$ (where $i = 0$ if $n = 2$ and $i = 3$ otherwise). Thus each instance of (S3) is true in the model, while (T5) is false. Hence (Def′), (S2), (S3), (S6), and (T1)–(T4) do not entail (T5).

Whichever way we link three- and four-termed comparative similarity, and whatever reasonable independent assumptions about similarity we make, (S3) does not entail the symmetry of similarity. Given (Def′), (S5) (and therefore (S1)) does entail symmetry; given (Def), it does not (for this reason, Lewis should not be read as tacitly relying on (Def′), since he accepts (S1) and rejects symmetry). The only unambiguous statement of symmetry is in $L(T)$. If we do see instances of (S3) as expressing symmetry, it is because we start with the four-termed relation and notice entailments in the other direction: given (Def) or (Def′), (T1)–(T5) entail (S3) but (T1)–(T4) do not. That is further reason to take the four-termed relation, and therefore degrees of similarity, as fundamental.

Appendix: Proofs For convenience, the axioms of the various theories mentioned in the results below are summarized here:

(Sa): (S1),(S2),(S3)	(Ta): (T1),(T2),(T3),(T4),(T5)
(Sb): (S2),(S3),(S5),(S6)	(Tb): (T2),(T3),(T4),(T5)
(Sc): (S1),(S3),(S7)	(Tc): (T1),(T2),(T3),(T5),(T6)
(Sd): (S3),(S6),(S7)	(Td): (T2),(T3),(T5),(T6)
	(Te): (T1),(T2),(T3),(T4),(T5),(T8)
(Sf): (S1),(S2),(S4)	(Tf): (T1),(T2),(T3),(T4)

Theorem I *(Sa) and (Ta) correspond.*

Proof: (i) We first show that (Ta) corresponds to (Sa). By the main text, it is enough to prove that (Ta) ∪ {(Def)} logically entails (Sa). Assuming (Def), (T1) gives (S1) and (T4) gives (S2). Each instance of (S3) requires repeated application of (T2) and (T5), following the pattern in Section 1.

(ii) We must now show that (Sa) corresponds to (Ta). Let (X,S) be a model of (Sa). We must define a relation T on X such that (X,S,T) is an $L(S,T)$-model of (Ta) ∪ {(Def)}. To this end, define a relation Q on X^2 by: $Q((w,x),(y,z))$ iff either $w = x$ or $(w,x) = (y,z)$ or both $x = y$ and $S(w,x,z)$. Q is obviously reflexive. Let Q′ be its ancestral; thus Q′ is reflexive and transitive. Define $\|$ by: $(w,x)\|(y,z)$ iff $Q'((w,x),(y,z))$ and $Q'((y,z),(w,x))$. Thus $\|$ is an equivalence relation. Let $\|w,x\|$ be the $\|$-equivalence class of (w,x). It is easily checked that the following definition is consistent: $Q''(\|w,x\|,\|y,z\|)$ iff $Q'((w,x),(y,z))$. Thus Q″ is reflexive, transitive, and antisymmetric. Now it is well known that every reflexive partial ordering of a set can be extended to a reflexive total ordering. Hence there is a reflexive, transitive, antisymmetric, connected relation $\leq$ such that $Q''(\|w,x\|,\|y,z\|)$ entails $\|w,x\| \leq \|y,z\|$. T can now be defined by: $T(w,x,y,z)$ iff $\|w,x\| \leq \|y,z\|$. We must show that (X,S,T) is a model of (Ta) ∪ {(Def)}.

The connectivity and transitivity of $\leq$ ensure that (T1) and (T2) respectively are true in (X,S,T). By definition we have $Q((x,x),(y,z))$ (for x,y,z in X), therefore $Q'((x,x),(y,z))$, therefore $Q''(\|x,x\|,\|y,z\|)$, therefore $\|x,x\| \leq \|y,z\|$; thus (T3) is true in (X,S,T). Now (X,S) is an $L(S)$-model of (S1), which entails that $S(x,y,x)$ for x and y in X, which gives us $Q((x,y),(y,x))$; arguing as before, $\|x,y\| \leq \|y,x\|$; thus (T5) is true in (X,S,T). Now observe that if $Q((w,x),(y,y))$ then $w = x$, for if $x = y$ and $S(w,x,y)$, $w = x$ follows because (X,S) is a model of (S2). Hence, by induction, if $Q'((w,x),(y,y))$ then $w = x$. Now suppose that $T(x,y,y,y)$. We have already seen that (X,S,T) is a model of (T3), so $T(y,y,x,y)$. Thus $\|x,y\| \leq \|y,y\|$ and $\|y,y\| \leq \|x,y\|$; since $\leq$ is antisymmetric, $\|x,y\| = \|y,y\|$, so $(x,y)\|(y,y)$, so $Q'((x,y),(y,y))$, so by the above $x = y$. Thus (T4) is true in (X,S,T).

It remains to show that (Def) is true in (X,S,T). One half is trivial. For suppose that $S(x,y,z)$. Then $Q((x,y),(y,z))$ and, since (X,S) is a model of (S1), $Q((y,z),(z,y))$; thus $Q'((x,y),(z,y))$, so $T(x,y,z,y)$. Conversely, suppose that $T(x,y,z,y)$. Suppose also, for a *reductio*, that $S(x,y,z)$ fails: so, by (S1) again, $S(z,y,x)$, so, by what we have just shown $T(z,y,x,y)$. Thus $\|x,y\| \leq \|z,y\|$ and $\|z,y\| \leq \|x,y\|$, so $\|x,y\| = \|z,y\|$, so $Q'((x,y),(z,y))$. By (S1), $Q'((y,x),(z,y))$. Since $S(x,y,x)$ without $S(x,y,z)$, $x \neq z$. Hence, there is a sequence $(y,x) =$

$(u_0,v_0),\ldots,(u_n,v_n) = (z,y)$, where for $i = 0,\ldots,n-1$, $Q((u_i,v_i), (u_{i+1},v_{i+1}))$, and we can assume that $(u_i,v_i) \neq (u_{i+1},v_{i+1})$. Hence, for each i, either $u_i = v_i$ or $u_{i+1} = v_i$ and $S(u_i,u_{i+1},v_{i+1})$. Now suppose that for some i, $u_{i+1} = v_{i+1}$ but $u_i \neq v_i$. Then $u_{i+1} = v_i$ and $S(u_i,u_{i+1},v_{i+1})$, so $S(u_i,u_{i+1},u_{i+1})$, so by (S2) $u_i = u_{i+1}$, so $u_i = v_i$, which is a contradiction. Hence if $u_i = v_i$ for some i, $u_0 = v_0$, that is, $x = y$. But $S(z,y,x)$, so if $x = y$ then $x = z$, which is impossible. Hence $x \neq y$, so for each i $u_i \neq v_i$, and thus $u_{i+1} = v_i$ and $S(u_i,u_{i+1},v_{i+1})$. Thus for $i = 0,\ldots,n-2$, $S(u_i,u_{i+1},u_{i+2})$. Moreover, $S(u_{n-1},u_n,u_0)$ because $S(u_{n-1}, u_n,v_n)$. Since (X,S) is a model of $(S3_n)$, whose antecedent we have just obtained, $S(u_1,u_0,u_n)$, that is, $S(x,y,z)$, contrary to hypothesis. Hence $T(x,y,z,y)$ entails $S(x,y,z)$. Thus (Def) is true in (X,S,T).

Theorem I′ *Let (X,S) be a model of (Sa). There is a unique relation T on X such that (X,S,T) is a model of $(Ta) \cup \{(Def)\}$ iff Q' (as defined in the proof of Theorem I) is connected on X^2.*

Proof: Notation from the proof of Theorem I is used throughout.

(i) Suppose that Q' is not connected. Then the relation Q'' on the $\|$-equivalence classes is not connected, and thus can be extended to distinct total orders $\leq$ and $\leq'$. By the proof of Theorem I, they give rise to distinct relations T and T′ on X such that both (X,S,T) and (X,S,T′) are models of (Ta) $\cup\{$(Def)$\}$.

(ii) Suppose that Q' is connected. We show that if (X,S,T) is a model of (Ta) $\cup$ $\{$(Def)$\}$, $T(w,x,y,z)$ iff $Q'((w,x),(y,z))$. Note that $Q((w,x),(y,z))$ implies $T(w,x,y,z)$: by (T3) if $w = x$, by (T7) if $(w,x) = (y,z)$ and by (T2), (T5), and (Def) if $x = y$ and $S(w,x,z)$. Thus by (T2) $Q'((w,x),(y,z))$ implies $T(w,x,y,z)$. Conversely, suppose for a *reductio* that $T(w,x,y,z)$ without $Q'((w,x),(y,z))$. Since Q' is connected, $Q'((y,z),(w,x))$. Hence there are pairs $(y,z) = (u_0,v_0),\ldots,(u_n,v_n) = (w,x)$ such that for $i = 0,\ldots,n-1$ $Q((u_i,v_i),(u_{i+1},v_{i+1}))$. We can assume that $(u_i,v_i) \neq (u_{i+1},v_{i+1})$. Suppose that $u_i = v_i$ for some i; then, as in the proof of Theorem I, $u_0 = v_0$, so $y = z$; but $T(w,x,y,z)$, so $T(w,x,x,x)$ would follow by (T2) and (T3), hence $w = x$ by (T4), hence $Q((w,x),(y,z))$, contrary to the hypothesis that $Q'((w,x),(y,z))$ fails. Thus for each i $u_i \neq v_i$, so $v_i = u_{i+1}$ and $S(u_i,v_i,v_{i+1})$. Now $Q'((y,z),(u_i,v_i))$ and $Q'((u_{i+1},v_{i+1}),(w,x))$, so, by what we have already shown, $T(y,z,u_i,v_i)$ and $T(u_{i+1},v_{i+1},w,x)$. But $T(w,x,y,z)$, so by (T2) $T(u_{i+1},v_{i+1},u_i,v_i)$. Since $v_i = u_{i+1}$, $T(v_i,v_{i+1},u_i,v_i)$, so by (T2), (T5), and (Def), $S(v_{i+1},v_i,u_i)$, so $Q((v_{i+1},u_{i+1}),(v_i,u_i))$. As in the proof of Theorem I, $Q((u_{i+1},v_{i+1}),(v_{i+1}, u_{i+1}))$ and $Q((v_i,u_i),(u_i,v_i))$, so $Q'((u_{i+1},v_{i+1}),(u_i,v_i))$. Since i was arbitrary, an induction gives $Q'((u_n,v_n),(u_0,v_0))$, i.e., $Q'((w,x),(y,z))$, which is a contradiction.

Theorem II (III, IV) *(Sb) and (Tb) ((Sc) and (Tc), (Sd) and (Td)) correspond.*

Proof: Similar to that of Theorem I.

Theorem II′ (III′, IV′) *Let (X,S) be a model of (Sb) ((Sc), (Sd)). There is a unique relation T on X such that (X,S,T) is a model of $(Tb) \cup \{(Def)\}$ ($(Tc) \cup \{(Def)\}$, $(Td) \cup \{(Def)\}$) iff Q' is connected on X^2.*

Proof: Similar to the proof of Theorem I′.

Theorem V *If B is a theory in $L(T)$ axiomatized by purely universal sentences, then $S(B)$ and B correspond.*

Proof: By definition, $B \cup \{(\mathrm{Def})\}$ entails $S(B)$, so B corresponds to $S(B)$. What we must show is that $S(B)$ corresponds to B. Let (X,S) be a model of $S(B)$. Introduce names for all the elements of X. Let D(X,S) be the set of all sentences true in the model (X,S)′ which results from adding these assignments of names to objects of (X,S) (the *complete diagram* of (X,S)′). Now suppose that $D(\mathrm{X,S}) \cup B \cup \{(\mathrm{Def})\}$ is inconsistent. Then some finite subset of it is inconsistent, so for some sentence θ in D(X,S), $B \cup \{(\mathrm{Def})\} \vdash \sim\theta$. Now the names in $\sim\theta$ do not appear in $B \cup \{(\mathrm{Def})\}$, and can thus be treated like variables. That is, if $Q(\sim\theta)$ is the result of replacing the names in $\sim\theta$ by distinct variables which do not appear in $\sim\theta$ and prefixing universal quantifiers to bind them, $B \cup \{(\mathrm{Def})\} \vdash Q(\sim\theta)$. But $Q(\sim\theta)$ is in $L(S)$ and so in $S(B)$. Thus $Q(\sim\theta)$ is true in (X,S)′; since $Q(\sim\theta) \vdash \sim\theta$, $\sim\theta$ is true in (X,S)′. But since θ is in D(X,S), θ is true in (X,S)′, which is a contradiction. Thus $D(\mathrm{X,S}) \cup B \cup \{(\mathrm{Def})\}$ is consistent, so it has a model M. By the construction of D(X,S), M has a submodel (X′,S′,T′) such that (X′,S′) is isomorphic to (X,S), where X′ is the set of elements of the domain of M to which names have been assigned. Since B is axiomatized by universal sentences and (Def) is a universal formula, (X′,S′,T′) is a model of $B \cup \{(\mathrm{Def})\}$. Hence there is a relation T on X such that $B \cup \{(\mathrm{Def})\}$ is true in (X,S,T).

Theorem VI *There is no theory A in $L(S)$ such that A and (Te) correspond.*

Proof: Suppose, for a contradiction, there there is a theory A in $L(S)$ such that A and (Te) correspond. Let (X,s) be a standard model of Peano Arithmetic (where s is the successor operation) and (X′,s') be a nonstandard model elementarily equivalent to (X,s). We define a real-valued function on X'^2 as follows:

$d(i,j) = 0$ if $i = j$
$d(i,j) = 2$ if i is standard in (X′,s') and either $s'(i) = j$ or $s'(j) = i$
$d(i,j) = 3$ if i is nonstandard in (X′,s') and either $s'(j) = i$ or $s'(i) = j$
$d(i,j) = 1$ otherwise.

Now define T′ by: $\mathrm{T}'(h,i,j,k)$ iff $d(h,i) \leq d(j,k)$. Since d obeys conditions (M1)–(M3) on a metric, (X′,T′) is a model of (T1)–(T5). Moreover, (X′,T′) is a model of (T8), for if h is a standard number in (X′,s') and i is nonstandard, there is no element k of X′ such that $\mathrm{T}'(i,s'(i),h,k)$. Since (Te) corresponds to A, there is a relation S′ on X′ such that (X′,S′,T′) is a model of $A \cup \{(\mathrm{Def})\}$. Clearly, $\mathrm{S}'(i,j,k)$ iff $d(i,j) \leq d(k,j)$. It can be routinely checked, then, that $\mathrm{S}'(i,j,k)$ iff either (a) $i = j$ or (b) $j \neq k$, $s'(i) \neq j$ and $s'(j) \neq i$ or (c) either $s'(j) = k$ or $s'(k) = j$. Thus S′ is the extension of a first-order open formula in the language of Peano Arithmetic in (X′,s'). Let S be the extension of this formula in (X,s). Since (X,s) and (X′,s') are elementarily equivalent, (X,S) and (X′,S′) are elementarily equivalent. But (X′,S′) is a model of A, so (X,S) is a model of A. Since A corresponds to (Te) there is a relation T on X such that (X,S,T) is a model of (Te) $\cup \{(\mathrm{Def})\}$. Now $\mathrm{S}(i,j,k)$ iff (a) $i = j$ or (b) $j \neq k$, $s(i) \neq j$ and $s(j) \neq i$ or (c) either $s(j) = k$ or $s(k) = j$. Thus for any i in X we have $\mathrm{S}(i,s(i),s(s(i)))$ and $\mathrm{S}(s(s(i)),s(i),i)$. Since (X,S,T) is a model of (Def), we have $T(i,s(i),s(s(i)),s(i))$ and $T(s(s(i)),s(i),i,s(i))$; since (T2) and (T5) are true in (X,S,T), this gives $T(i,s(i),s(i),s(s(i)))$ and $T(s(i),s(s(i)),i,s(i))$. Since (X,s) is a standard model of arithmetic, for any i and j in X either we can reach i from j by a finite number of applications of s or vice versa. In either case, by repeated application of (T2) to the last result, we have $T(i,s(i),j,s(j))$. Now for any h in X we have $S(h,i,s(i))$, and therefore $T(h,i,s(i),i)$ by (Def), and hence $T(h,i,i,s(i))$ by (T2) and (T5). By (T2), $T(h,i,j,s(j))$ for any h, i, and j in X. But this is a contradiction, since (T8) was supposed to be true in (X,S,T).

Theorem VII *(Sa) is not finitely axiomatizable.*

Proof: Suppose, for a contradiction, that (Sa) is finitely axiomatizable in $L(S)$. Hence there is a sentence ϕ (the conjunction of the finite number of axioms) in $L(S)$ of which (S1)–(S2) and each ($S3_i$) is a logical consequence; moreover ϕ is a logical consequence of (S1)–(S2) and the ($S3_i$)'s, so that ϕ is a logical consequence of (S1)–(S2) and some finite subset of the ($S3_i$)'s. Thus each of the ($S3_i$)'s is a logical consequence of some finite subset of them and (S1)–(S2). Hence it suffices for the proof to find, for each k, a model of (S1)–(S2) and ($S3_2$)–($S3_k$) in which some ($S3_j$) is false.

Let R be any asymmetric many-one relation on a set X. Thus $R(x,y)$ excludes $R(y,x)$ and if $R(x,y)$ and $R(x,z)$, $y = z$. Define a real-valued function d on pairs of elements of X as follows:

$d(x,y) = 0$ if $x = y$
$d(x,y) = 1$ if $R(x,y)$
$d(x,y) = 2$ if $R(y,x)$
$d(x,y) = 3$ otherwise.

Note that $d(x,y) = 0$ iff $d(y,x) = 0$; $d(x,y) = 1$ iff $d(y,x) = 2$; $d(x,y) = 3$ iff $d(y,x) = 3$. Now define a relation S on X by: $S(x,y,z)$ iff $d(x,y) \leq d(z,y)$. By the connectivity of $\leq$, (X,S) is a model of (S1); since d obeys conditions (M1)–(M2) on a metric, (X,S) is a model of (S2). Now define an (R,n)-*chain* to be a sequence $(x_0, \ldots, x_n)$ of elements of X such that $R(x_i,x_{i+1})$ for $i < n$. An (R,n)-*cycle* is a sequence $(x_0, \ldots, x_n)$ such that $(x_0, \ldots, x_n, x_0)$ is an (R,$n + 1$)-chain. $\oplus$ is addition modulo $n + 1$ (thus $n \oplus 1 = 0$). We now show that (X,S) is a model of ($S3_2$)–($S3_k$) iff there is no (R,n)-cycle for $n \leq k$.

Suppose that $(v_0, \ldots, v_n)$ is an (R,n)-cycle. Hence for $i < n$, $d(v_i, v_{i\oplus 1}) = 1$ and $d(v_{i\oplus 2}, v_{i\oplus 1}) = 2$, so $S(v_i, v_{i\oplus 1}, v_{i\oplus 2})$. However, $d(v_1, v_0) = 2$ and $d(v_n, v_0) = 1$, so $S(v_1, v_0, v_n)$ fails. Then (X,S) would not be a model of ($S3_n$). Hence if (X,S) is a model of ($S3_2$)–($S3_k$), there is no (R,n)-cycle for $n \leq k$.

Conversely, suppose that (X,S) is not a model of ($S3_n$), where $n \leq k$. Thus, for some $v_0, \ldots, v_n$, $S(v_i, v_{i\oplus 1}, v_{i\oplus 2})$ for $i < n$, while $S(v_1, v_0, v_n)$ fails. Since (X,S) is a model of (S1), we have $S(v_n, v_0, v_1)$ instead. Thus $S(v_i, v_{i\oplus 1}, v_{i\oplus 2})$ for $i \leq n$: that is, $d(v_i, v_{i\oplus 1}) \leq d(v_{i\oplus 2}, v_{i\oplus 1})$. Suppose that, for some i, $d(v_i, v_{i\oplus 1}) = 0$, where $0 < i$; since $d(v_{i-1}, v_i) \leq d(v_{i\oplus 1}, v_i)$, $d(v_{i-1}, v_i) = 0$; thus, by induction, $d(v_0, v_1) = 0$, so $d(v_1, v_0) \leq d(v_n, v_0)$, which is impossible, since $S(v_1, v_0, v_n)$ fails. Thus, for each i, $1 \leq d(v_i, v_{i\oplus 1})$. Similarly, suppose that, for some i, $d(v_i, v_{i\oplus 1}) = 3$; since $d(v_i, v_{i\oplus 1}) \leq d(v_{i\oplus 2}, v_{i\oplus 1})$, $d(v_{i\oplus 1}, v_{i\oplus 2}) =$

$d(v_{i\oplus 2}, v_{i\oplus 1}) = 3$; thus, by induction $d(v_n, v_0) = 3$, which is impossible, since $S(v_1, v_0, v_n)$ fails. Thus, for each i, $d(v_i, v_{i\oplus 1}) \leq 2$. Now suppose, for a contradiction, that there is no (R,j)-cycle for $j \leq n$. Since $S(v_1, v_0, v_n)$ fails, $1 \leq d(v_n, v_0) < d(v_1, v_0) \leq 2$. Thus $d(v_n, v_0) = 1$ and $d(v_1, v_0) = 2$, so $R(v_n, v_0)$ and $R(v_0, v_1)$. Note that (v_0, v_1) is an (R,1)-chain from v_0 to v_1. Let h be the greatest value of i such that, for some j, there is an (R,j)-chain from v_0 to v_i with $j \leq i \leq n$. Thus there is an (R,j)-chain $(u_0, \ldots, u_j)$, where $j \leq h$, $u_0 = v_0$, $u_j = v_h$ and $R(u_{j-1}, v_h)$. Suppose that $h < n$. Thus $h < h \oplus 1$. If $R(v_h, v_{h\oplus 1})$, then $(u_0, \ldots, u_j, v_{h\oplus 1})$ is an (R,$j \oplus 1$)-chain from v_0 to $v_{h\oplus 1}$, where $j \oplus 1 \leq h \oplus 1$, which contradicts the definition of h. Thus $R(v_h, v_{h\oplus 1})$ fails. Hence $d(v_h, v_{h\oplus 1}) \neq 1$, so $d(v_h, v_{h\oplus 1}) = 2$, so $R(v_{h\oplus 1}, v_h)$. Hence $2 = d(v_h, v_{h\oplus 1}) \leq d(v_{h\oplus 2}, v_{h\oplus 1}) \leq 2$, so $d(v_{h\oplus 2}, v_{h\oplus 1}) = 2$, so $R(v_{h\oplus 1}, v_{h\oplus 2})$. Since $R(v_{h\oplus 1}, v_h)$ and R is many-one, $v_{h\oplus 2} = v_h$. If $h \oplus 2 = 0$, $v_h = v_0$, so $R(u_{j-1}, u_0)$ (since $u_0 = v_0$ and $R(u_{j-1}, v_h)$); but then $(u_0, \ldots, u_{j-1})$ would be an (R,$j-1$)-cycle, with $j - 1 < h \leq n$, contrary to hypothesis. Thus $h \oplus 2 \neq 0$, so $h < h \oplus 2$. But since $u_j = v_h = v_{h\oplus 2}$, $(u_0, \ldots, u_j)$ is an (R,j)-chain from v_0 to $v_{h\oplus 2}$, where $j \leq h \oplus 2$, which contradicts the definition of h. Thus the assumption that $h < n$ has led to contradiction, so $h = n$. Hence $(u_0, \ldots, u_j)$ is an (R,j)-chain from v_0 to v_n. But we have already proved that $R(v_n, v_0)$, so $(u_0, \ldots, u_j)$ is an (R,j)-cycle, with $j \leq n$, contradicting the assumption that there are no such cycles. Hence if (X,S) is not a model of $(S3_2)$–$(S3_k)$, there is an (R,j)-cycle, where $j \leq k$.

Thus, by the previous two paragraphs, (X,S) is a model of $(S3_2)$–$(S3_n)$ iff there is no (R,j)-cycle for $j \leq n$. Where $2 \leq n$, let X be the set of integers $\{0, \ldots, n\}$ and $\oplus$ the operation of addition modulo $n + 1$ on them. Define R by: $R(x,y)$ iff $x \oplus 1 = y$. R obeys the required conditions: if $x \oplus 1 = y$ then $y \oplus 1 \neq x$; if $x \oplus 1 = z$ and $y \oplus 1 = z$ then $x = y$. Thus $(0, \ldots, n)$ is an (R,n)-cycle, but there is no (R,j)-cycle for $j < n$. Define S in terms of R as above. Thus (X,S) is a model of (S1)–(S2) and $(S3_2)$–$(S3_{n-1})$ but not of $(S3_n)$. Since n can be taken to be arbitrarily large, this is what we required.

Corollary *(Sb), (Sc), and (Sd) are not finitely axiomatizable.*

Proof: Each of these theories includes all instances of (S3) and is included in (Sa). Hence if one of them could be finitely axiomatized, (Sa) could be finitely axiomatized by the addition as axioms of whichever of (S1)–(S2) were missing.

Theorem VIII *(Sf) and (Tf) correspond.*

Proof: Similar to that of Theorem I.

Independence results: To prove that (T1) is independent of (T2)–(T7), let the variables range over the subsets of a two-membered set, let $f(x,y)$ be the symmetric difference of x and y and $T(w,x,y,z)$ be true iff $f(w,x) \subseteq f(y,z)$. On this interpretation (T1) is false while T(2)–T(7) are true. Similarly, by (Def), (S1) is independent of (S2)–(S7).

To prove that (T2) is independent of (T1) and (T3)–(T7), let the variables range over real numbers and $T(w,x,y,z)$ be true iff either $w = x$ or both $y \neq z$ and $(w - x)^2 < (y - z)^2 + 6$. Then (T1) and (T3)–(T7) are true but (T2) is false, since we have $T(3,0,2,0)$ and $T(2,0,1,0)$ without $T(3,0,1,0)$.

To prove that (T3) is independent of (T1)–(T2) and (T4)–(T7), let the variables range over the real numbers, define d by:

$$d(x,y) = -(x+y)^2 \text{ if } x = y$$
$$d(x,y) = (x-y)^2 \text{ otherwise,}$$

and let $T(w,x,y,z)$ be true iff $d(w,x) \leq d(y,z)$. On this interpretation (T3) is false (since $T(0,0,1,0)$ fails) while (T1)–(T2) and (T4)–(T7) are true.

To prove that (T4) is independent of (T1)–(T3) and (T5)–(T7), let R be any equivalence relation other than identity and let $T(w,x,y,z)$ be true iff either $R(w,x)$ is true or $R(y,z)$ is false; similarly (by (Def)), (S2) is independent of (S1) and (S3)–(S7).

To prove that (T5) is independent of (T1)–(T4) and (T6)–(T7), note that by previous results, S((Tf)) is finitely axiomatizable while S((Ta)) is not.

To prove that (T6) is independent of (T1)–(T3), (T5), and (T7), let the variables range over the real numbers, define d by:

$$d(x,y) = 0 \text{ if } (x-y)^2 \leq 1$$
$$d(x,y) = 1 \text{ otherwise,}$$

and let $T(w,x,y,z)$ be true iff $d(w,x) \leq d(y,z)$. Then (T1)–(T3), (T5), and (T7) are true but (T6) is false (since $T(0,1,1,1)$ is true and $T(0,2,1,2)$ is false). Similarly, (S7) is independent of (S1) and (S3)–(S6).

To prove that (S5) is independent of (S2)–(S4) and (S6), let $S(x,y,z)$ be true iff $x = y$.

To prove that (S6) is independent of (S2)–(S5) and (S7), let $S(x,y,z)$ be true iff $d(x,y) = d(z,y)$ for a nontrivial metric d.

Finally, note that for each instance of (S3), the proof of Theorem VII gives a model of (S1), (S2), and (S4)–(S7) in which it is false.

NOTES

1. *Proof:* Their claim entails that there are at least as many degrees of similarity as cardinals greater than c, but there are more cardinals greater than c than real numbers.

2. Tversky has given doubts about the symmetry of similarity an empirical basis ([18], pp. 328, 333–338). His experiments confirm the hypothesis that when we assess the similarity of a to b, we give more weight to salient features of a that b lacks than to features of b that a lacks: 'the less salient stimulus is more similar to the salient stimulus than vice versa' (ibid. p. 333); a comparable effect occurs if the perceived similarity of a to b is measured by the probability of one being mistaken for the other. On the other hand, if our task is phrased as "assess the similarity of a and b", we are likely to use a symmetrical method. Of course, the experiments could be interpreted just as well in terms of a subjective bias in similarity judgments (perhaps to be explained by differential attention) as in terms of a nonsymmetry in similarity; in themselves they are neutral.

3. (S1), (S4), and (S2) (when restricted to possible worlds) are Lewis's assumptions (1), (2), and (4) about similarity respectively; his other assumptions concern the relation of accessibility between worlds and thus are not relevant to a general logic of similarity. (S3) is the reflexive counterpart of the schema discussed in Section 1; in effect,

Lewis rejected ($S3_2$). (S1) is also in effect clause I of the definition of a *strongly conditional absolute difference structure* at [10], p. 183; (S2) is clause 2; ($S3_2$) and ($S3_3$) are 3(ii) and 3(iii), respectively; (S4) is 3(i); the other clauses depend on the assumption of a unidimensional domain.

4. Cf. also [17]; [3], pp. 130–132 is an example of the use of nonuniversal axioms for comparative similarity and has an interesting discussion of their legitimacy, cf. [1] and Sections 6.6 and 9.5 of [14]; the definitions of absolute-difference structures and strongly conditional absolute-difference structures at pp. 172 and 183 of [10], respectively, both involve existential clauses.

REFERENCES

[1] Adams, E. W., R. F. Fagot, and R. E. Robinson, "On the empirical status of axioms in theories of fundamental measurement," *Journal of Mathematical Psychology*, vol. 7 (1970), pp. 379–409.

[2] Arrow, K. J., "Values and collective decision-making," pp. 110–126 in *Philosophy and Economic Theory*, eds. F. Hahn and M. Hollis, Oxford University Press, Oxford, 1979.

[3] Beals, R., D. H. Krantz, and A. Tversky, "Foundations of multidimensional scaling," *Psychological Review*, vol. 75 (1968), pp. 127–142.

[4] Bugajski, S., "Languages of similarity," *Journal of Philosophical Logic*, vol. 12 (1983), pp. 1–18.

[5] Davies, M., *Meaning, Quantification, Necessity*, Routledge, London, 1981.

[6] Forrest, P. and D. M. Armstrong, "An argument against David Lewis' theory of possible worlds," *Australasian Journal of Philosophy*, vol. 62 (1984), pp. 164–168.

[7] Goodman, N., "Seven strictures on similarity," pp. 19–29 in *Experience and Theory*, eds. L. Foster and J. W. Swanson, University of Massachusetts Press, Boston, 1970.

[8] Hilpinen, R., "Approximate truth and truthlikeness," pp. 19–42 in *Formal Methods in the Methodology of Empirical Sciences*, eds. M. Przelecki, K. Szaniawski, and R. Wojcicki, Reidel, Dordrecht, 1976.

[9] James, W., *The Principles of Psychology*, Holt, New York, 1890.

[10] Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky, *Foundations of Measurement*, Vol. I, Academic Press, New York, 1971.

[11] Lewis, D. K., *Counterfactuals*, Blackwell, Oxford, 1973.

[12] Lewis, D. K., *On the Plurality of Worlds*, Blackwell, Oxford, 1986.

[13] Nicod, J., *Geometry and Induction*, Routledge, London, 1970.

[14] Pfanzagl, J., *Theory of Measurement*, Wiley, New York, 1968.

[15] Quine, W. V., *Ontological Relativity and Other Essays*, Columbia University Press, New York, 1969.

[16] Quine, W. V., *The Roots of Reference*, Open Court, La Salle, 1973.

[17] Scott, D. and P. Suppes, "Foundational Aspects of Theories of Measurement," *The Journal of Symbolic Logic*, vol. 23 (1958), pp. 113–128.

[18] Tversky, A., "Features of Similarity," *Psychological Review*, vol. 84 (1977), pp. 327–352.

[19] Williamson, T., "The infinite commitment of finite minds," *Canadian Journal of Philosophy*, vol. 14 (1984), pp. 235–255.

Department of Philosophy
Trinity College
University of Dublin
Dublin, Ireland