THE JOURNAL OF PHILOSOPHY

VOLUME CVII, NUMBER 9
SEPTEMBER 2010

Published by **The Journal of Philosophy, Inc.**

THE JOURNAL OF PHILOSOPHY

FOUNDED BY FREDERICK J. E. WOODBRIDGE AND WENDELL T. BUSH

Purpose: To publish philosophical articles of current interest and encourage the interchange of ideas, especially the exploration of the borderline between philosophy and other disciplines.

All communications to the Editors and Trustees and all manuscripts may be sent to Alyssa Timin, Managing Editor, at submissions@journalofphilosophy.org.

You may also visit our website at: www.journalofphilosophy.org

THE JOURNAL OF PHILOSOPHY

2010

SUBSCRIPTIONS (12 issues)

Individuals	$45.00
Libraries and Institutions	$100.00
Libraries and Institutions (print and online)	$180.00
Students, retired/unemployed philosophers	$20.00
Postage outside the U.S.	$15.00

Payments only in U.S. currency on a U.S. bank. All back volumes and separate issues available back to 1904. Please inquire for advertising rates; ad space is limited, so ad reservations are required.

Published monthly as of January 1977; typeset by Dartmouth Journal Services, Waterbury, VT, and printed by The Sheridan Press, Hanover, PA.

All communication about subscriptions and advertisements may be sent to Pamela Ward, Business Manager, Mail Code 4972, 1150 Amsterdam Avenue, Columbia University, New York, NY 10027. (212) 866-1742

The Journal of Philosophy (ISSN 0022-362X) is published monthly by The Journal of Philosophy, Inc., MC4972, Columbia University, 1150 Amsterdam Avenue, New York, NY 10027. Periodicals postage paid at New York, NY, and other mailing offices.

POSTMASTER: Please send address changes to *The Journal of Philosophy* at MC4972, Columbia University, 1150 Amsterdam Avenue, New York, NY 10027.

ISSN 0022–362X

THE JOURNAL OF PHILOSOPHY
VOLUME CVII, NO. 9, SEPTEMBER 2010

LAMBDA IN SENTENCES WITH DESIGNATORS: AN ODE TO COMPLEX PREDICATION*

In his classic studies of the "Logic of Sense and Denotation" (*LSD*), Alonzo Church contrasted three criteria or explications of the notion of *sense*, or semantic content, and of the resulting notion of synonymy, in the sense of *having the same sense*.[1] These are Alternatives (0), (1), and (2), numbered in order from the most strict to the least. On the most lax criterion, Alternative (2), synonymy is taken to be mere logical equivalence. On the intermediate Alternative (1), the criterion for synonymy is convertibility by means of Church's

*I am grateful to the participants in my seminars at the University of California, Santa Barbara during Winter 2007 and at the City University of New York Graduate Center during Fall 2009 and to my audiences at the University of Missouri Kline Workshop of Meaning, April 2009 and at the University of Buffalo, October 2009 for discussion of Kripke's critique and my response. I thank Luke Manning and Teresa Robertson for comments and suggestions. I am especially grateful to C. Anthony Anderson for valuable suggestions and for valuable discussion, both of the relevant issues and of the relevant and brilliant work of our remarkable former teacher, the late Alonzo Church.

[1] Church, "Abstract of 'A Formulation of the Logic of Sense and Denotation'," *Journal of Symbolic Logic*, XXI, 1 (March 1946): 31; "A Formulation of the Logic of Sense and Denotation," in Paul Henle, Horace Kallen, and Susanne Langer, eds., *Structure, Method and Meaning: Essays in Honor of Henry M. Sheffer* (New York: Liberal Arts Press, 1951), pp. 3–24; "Outline of a Revised Formulation of the Logic of Sense and Denotation," Part 1, *Noûs*, VII, 1 (March 1973): 24–33 and Part 2, VIII, 2 (May 1974): 135–56; "A Revised Formulation of the Logic of Sense and Denotation, Alternative (1)," *Noûs*, XXVII, 2 (June 1993): 141–57.

For some subsequent illuminating work on Church's *LSD*, see C. Anthony Anderson, "Alternative (I*): A Criterion of Identity for Intensional Entities," in Anderson and Michael Zelëny, eds., *Logic, Meaning and Computation: Essays in Memory of Alonzo Church* (Boston: Kluwer, 2001), pp. 395–427, at pp. 421–22. There is a valuable discussion of *LSD* and Church's three alternative criteria for synonymy in Anderson's "Alonzo Church's Contributions to Philosophy and Intensional Logic," *The Bulletin of Symbolic Logic*, IV, 2 (June 1998): 129–71. I thank Anderson for bibliographical references. (It is untrue that a psychedelic Lennon-McCartney composition contains a veiled reference to Church's classic study.)

0022-362X/10/0709/445–468

λ-abstraction operator.[2] Expressions are *λ-convertible* if one is obtainable from the other by a sequence of applications of the λ-conversion rules of λ-expansion, λ-contraction, and alphabetic change of bound variables. Since the λ-conversion rules are each logically reversible (that is, the reverse inference is equally valid), expressions that are λ-convertible are deemed "synonymous" on both Alternatives (1) and (2). On the strictest criterion, Alternative (0), the criterion for synonymy is Church's notion of synonymous isomorphism, an improvement over Rudolf Carnap's notion of intensional isomorphism.[3] Expressions are *synonymously isomorphic* if one is obtainable from the other by a sequence of alphabetic changes of bound variables or replacements of component expressions by syntactically simple synonyms. In particular, on Alternative (0) the inference rules of λ-expansion and λ-contraction are not deemed to preserve sense. Aside from interchange of synonyms, at least one of which is syntactically simple,[4] the only interchange of logical equivalents deemed strictly to preserve sense is alphabetic change of bound variables. Logically equivalent sentences that are not synonymously isomorphic are deemed not *strictly synonymous*.

Few language philosophers today stubbornly maintain that logically equivalent sentences have the same semantic content, in the sense of expressing the same proposition.[5] In my judgment, those philosophers who do, confuse the semantic content of a sentence with a different semantic value, which I call the *logical content*.[6] Logically

[2] The λ-abstraction operator is a variable-binding operator. The extension, with respect to an assignment *s* of values to variables, of a λ-abstract $\ulcorner(\lambda\alpha)[\zeta_\alpha]\urcorner$ is the function that assigns to any potential value *v* in the range of the variable α, the extension of ζ_α with respect to the assignment that assigns *v* to α and is otherwise the same as *s*. If ζ_α is a singular term, then $\ulcorner(\lambda\alpha)[\zeta_\alpha]\urcorner$ is a compound functor. If ζ_α is a formula, then $\ulcorner(\lambda\alpha)[\zeta_\alpha]\urcorner$ is a compound predicate. The λ-conversion rule of *λ-expansion* licenses the replacement within a formula of any occurrence of ζ_β by an occurrence of $\ulcorner(\lambda\alpha)[\zeta_\alpha](\beta)\urcorner$, where β is of the same syntactic type as α and ζ_β is the result of uniformly substituting free occurrences of β for the free occurrences of α throughout ζ_α. The λ-conversion rule of *λ-contraction* licenses the reverse replacement of $\ulcorner(\lambda\alpha)[\zeta_\alpha](\beta)\urcorner$ by ζ_β.

[3] Church makes this proposal in his masterly essay, "Intensional Isomorphism and Identity of Belief," *Philosophical Studies*, v, 5 (October 1954): 65–73; reprinted in Nathan Salmon and Scott Soames, eds., *Propositions and Attitudes* (New York: Oxford, 1988), pp. 159–68.

[4] Interchange of syntactically compound expressions is not permitted. Church's idea seems to have been that there are no primitive synonymy rules stating that two compound expressions are synonymous. Instead, the synonymy of compound expressions is determined by the compositional semantic rules that govern the contents of compound expressions. (Thanks to C. Anthony Anderson for discussion.)

[5] Two prominent contemporary philosophers who maintain something close to an adherence to Church's Alternative (2) are Jaakko Hintikka and Robert Stalnaker. David Lewis's disciples also belong to this camp. In general, followers of Frege and Russell do not.

[6] This may be what Frege intended by his notion of *Erkenntniswerte* ("cognitive value"). Cf. my "On Content," *Mind*, CI, 404 (October 1992): 733–51; reprinted in my *Metaphysics, Mathematics, and Meaning: Philosophical Papers I* (New York: Oxford, 2005), pp. 269–85.

equivalent sentences are exactly those that share the same logical content. The logical content of a sentence might be identified with the class of its models. Sentences that are exactly alike in logical content might yet differ in semantic content, in the propositions they semantically express. One example of such pairs of sentences is the logical theorem '$(p \supset q) \vee (q \supset r)$' and Peirce's law, '$[(p \supset q) \supset p] \supset p$'. Another example might be 'Snow is white' and 'The number that is one if snow is white, and is zero if snow is not white, is one'. One example not involving logical truths is 'If a trespasser is caught, then he/she is prosecuted' and 'No trespassers are caught unless they are prosecuted'. Since these sentences are perfectly understandable, the fact that their equivalence is confirmed only upon reflection would seem to indicate that there is a difference in semantic content, although there is no difference in logical content.

Under careful scrutiny Alternative (1) fares little better than Alternative (2). Another pair of nonsynonymous logical equivalents is obtained from the sentence,

(A_E) This yacht is larger than that yacht is.

Suppose that, unbeknownst to the speaker, the occurrences of both complex demonstratives 'this yacht' and 'that yacht' are uttered with reference to the very same yacht. Even so, (A_E) contrasts sharply in semantic content with

(B_E) That yacht is a thing that is larger than it itself is.[7]

If the guest in Russell's famous example, on finally seeing the yacht, had compared it ("that yacht") to the one shown to him in a deceptive photograph ("this yacht"), he might well have believed the proposition expressed by (A_E) without thereby believing the proposition expressed by (B_E). This consideration is, by itself, already sufficient to prescribe contrasting semantic analyses of (A_E) and (B_E). The contrast between (A_E) and (B_E) also provides an explanation, or at least the beginning of an explanation, of how one can believe of the relevant yacht that it is larger than it is without believing that something is larger than itself. Distinguishing the contents of (A_E) and (B_E)

[7] I would formulate this as 'That yacht is a thing that is larger than oneself', except that use of the personal reflexive pronoun 'oneself' here, rather than the impersonal 'itself', is of questionable grammaticality. Still, the latter reduces the temptation to read the reflexive-pronoun occurrence as designating the yacht in question and instead encourages reading it as a bound variable. (I strongly suspect the former construal is incorrect. See the following note.) The English language does not present a happy alternative. As will soon become apparent, the reading I intend is best captured using a recurrent bound variable in lieu of a reflexive pronoun: 'That yacht is a thing x such that x is larger than x is'.

also provides the beginning of an explanation of why it is that one who believes the content of (A_E) is in no position to see that his belief is inconsistent. If (A_E) is read instead as expressing no more or less than what is expressed in (B_E), no such explanations are forthcoming.[8]

The differing semantic analyses of (A_E) and (B_E) reflect some of the structural differences between the sentences. Sentence (A_E) attributes a binary relation between a pair of objects—the mentioned yacht and that same yacht—whereas (B_E) attributes an impossible property to the yacht: being self-larger. The proposition expressed by (B_E)—the proposition about the yacht, that it is a thing-larger-than-it-itself—applies a notion of *reflexivity*.[9] The proposition expressed by (A_E) does not invoke reflexivity or anything else out of the ordinary.

The special semantic properties of (B_E) that distinguish it from (A_E) are captured by using the λ-operator:

(λx)[x is larger than x is](that yacht).

This is to be read:

That yacht is a thing x such that x is larger than x is.

If the complex demonstratives in (A_E) and (B_E) are now replaced by a single proper name '*a*' for the yacht in question, the resulting sentences respectively preserve the propositions expressed:

(A) a is larger than a is.
(B) (λx)[x is larger than x is](a).

These sentences therefore do not express the same proposition. The λ-abstract in (B) expresses a property or concept not expressed in (A): that of being self-larger. Yet (A) and (B) are logically equivalent, by the rules of λ-expansion (which licenses the inference from (A) to (B)) and λ-contraction (which licenses the reverse inference). This result discredits Alternative (1) as an analysis of propositional content.[10]

[8] Cf. Salmon, "Reflexivity," *Notre Dame Journal of Formal Logic*, XXVII, 3 (July 1986): 401–29, reprinted in Salmon and Soames, eds., *op. cit.*, pp. 240–74; "Reflections on Reflexivity," *Linguistics and Philosophy*, XV, 1 (February 1992): 53–63; "Illogical Belief," in James E. Tomberlin, ed., *Philosophical Perspectives 3: Philosophy of Mind and Action Theory* (Atascadero, CA: Ridgeview, 1989), pp. 243–85; and "The Resilience of Illogical Belief," *Noûs*, XL, 2 (June 2006): 369–75. All are reprinted in *Content, Cognition, and Communication: Philosophical Papers II* (New York: Oxford, 2007), pp. 30–64, 191–227. Cf. also my "Pronouns as Variables," *Philosophy and Phenomenological Research*, LXXII, 3 (May 2006): 656–64, reprinted in *Metaphysics, Mathematics, and Meaning: Philosophical Papers I*, pp. 399–406; and "Constraint with Restraint," in Gary Ostertag, ed., forthcoming festschrift for Stephen Schiffer.

[9] By 'reflexivity' I mean the notion of reflexivization expressed by '(λR)[(λx)[$R(x, x)$]]'.

[10] As a referee for this JOURNAL points out, there are examples of the same phenomenon from pure mathematics. One such example is '17 is evenly divisible by an

To reiterate: Where the two complex demonstratives occurring in (A_E) and the proper name '*a*' occurring in (A) all designate the same yacht, (A) expresses the same proposition as (A_E), (B) the same proposition as (B_E); yet (A) and (B) express different propositions.[11]

Saul Kripke is dubious.[12] He raises three considerations that disincline him to accept this account of the semantic content of

---

integer n iff $n = 17 \vee n = 1$' and '17 is prime'. One could come to believe what is expressed by the former without thereby believing that 17 is evenly divisible only by itself and 1. Therefore, different propositions are expressed.

[11] Kit Fine, in *Semantic Relationism* (Malden, MA: Blackwell, 2007), at p. 69, evidently misunderstands me as claiming that (A_E) and (A) semantically express different propositions while (A) and (B) express the same proposition—the reverse of my actual view. Fine defends and develops a view first proffered by Hilary Putnam in "Synonymy, and the Analysis of Belief Sentences," *Analysis*, XIV, 5 (1954): 114–22, reprinted in Salmon and Soames, eds., *op. cit.*, pp. 149–58, and later championed by David Kaplan in "Words," *Proceedings of the Aristotelian Society, Supplementary Volumes*, LXIV (1990): 93–119, at p. 95n6. The basic idea is that, where α and β are exactly synonymous terms (terms having the very same semantic content), $\phi_{\alpha\beta}$ is a sentence containing free occurrences of both terms, and $\phi_{\alpha\alpha}$ is the result of substituting free occurrences of α for free occurrences of β in $\phi_{\alpha\beta}$, the two sentences semantically expresses different propositions—as, for example, 'Bachelors socialize with other bachelors' and 'Unmarried men socialize with other bachelors' (assuming that 'bachelor' and 'unmarried man' are exactly synonymous). Putnam et al. contend that $\varphi_{\alpha\alpha}$ (at least normally) expresses a proposition that in some manner reflects additional material (additional structure, information, or something) that normally results from α's recurrence whereas $\phi_{\alpha\beta}$ does not. Kaplan and Fine maintain that (A_E) and (A) likewise (at least typically) express different propositions.

Church leveled a powerful criticism of the position of Putnam et al. in "Intensional Isomorphism and Identity of Belief": the position has the unwelcome consequence that the proposition expressed by the English sentence 'Unmarried men socialize with other bachelors' is inexpressible in a language that has only a phrase but no single word (or additional phrase) corresponding to the English 'unmarried man' (or else, at best, the proposition is expressible in such a language only by means of an allegedly ambiguous construction, and then only by means of an allegedly strained reading). To my knowledge, none of the view's adherents have addressed Church's observation that this consequence is excessively implausible. (The criticism, as presented here, involves minor extrapolation in conformity with Church's intent.)

In "Recurrence" (unpublished), I make further criticism of Fine's theory.

[12] Kripke, "Russell's Notion of Scope," *Mind*, CXIV, 456 (October 2005): 1005–37, at p. 1025n45.

I take this opportunity to correct Kripke's characterization at p. 1022 of our communication concerning Russell's example. In that discussion I emphasized the distinction in semantic content that I draw, and of which Kripke is dubious, between the binary-relational predication '*a* is larger than *a* is' and the monadic-predicational '*a* is a thing larger than itself'. I used the distinction not to solve a particular problem Kripke had noticed in Russell's discussion of his example, but rather to support my contention (which Kripke does not accept) that it is possible for one to believe, concerning a particular yacht *a*, that *a* is larger than *a* is while not thereby believing that *a* is self-larger (that is, a thing *x* larger than *x*). I was aware that this distinction (even if it is legitimate, as I maintain) does not solve the problem Kripke had noticed. For more on the relevance to Russell's example, see my "Points, Complexes, Complex Points,

λ-abstraction. Kripke's objections, as well as discussions I have had with him, indicate that he strongly favors Alternative (1) (or something close to it). The disagreement between us is pointed and fundamental. It is my considered judgment that Alternative (1) is incorrect and its advocacy a significant leap backward.

I

One of Kripke's objections is that my account of the semantic content of λ-abstraction threatens to call into question the validity of the standard proof in modal logic of the necessity of identity:

$$(x)(y)[x = y \supset \Box(x = y)].$$

Equally, Kripke argues, my account calls into question the validity of my own disproof of the popular thesis that, for some pairs of objects, there is no fact of the matter concerning whether they are one and the same or numerically distinct.[13] The necessity of identity is proved as follows: First, one proves a trivial lemma by applying the modal rule of necessitation to the reflexive law of identity, to obtain '$\Box(x = x)$'. The theorem is then easily proved by assuming '$x = y$', then invoking Leibniz's law of substitution to replace the second occurrence of 'x' in the lemma with 'y'. The disproof of the indeterminacy of identity invokes an application of Leibniz's law using an analogous lemma of the logic of determinacy: that there is a fact of the matter concerning whether x is x.[14] Kripke's objection to my account of the content of abstraction is the following:

> Someone might argue against the necessity of identity by claiming that only the self-identity of x is necessary, while the identity of x and x is contingent. Similarly, he or she might "refute" Salmon's own argument against vague [that is, indeterminate] identity by a parallel argument. Surely Salmon should be wary of this. I am.[15]

and a Yacht," in Nicholas Griffin and Dale Jacquette, eds., *Russell vs. Meinong: The Legacy of "On Denoting"* (New York: Routledge, 2008), pp. 343–63.

[13] See my *Reference and Essence* (Amherst, NY: Prometheus Books, 1981, 2005), at pp. 241–45; and "Identity Facts," in Christopher Hill, ed., *Philosophical Topics*, XXX, 1 (Spring 2002): 237–67, reprinted in *Metaphysics, Mathematics, and Meaning*, chapter 10.

[14] *Proof*: Assume for a *reductio* that there is a pair of objects, x and y, such that there is no fact concerning whether x and y are the same object. By the lemma, there is a fact concerning whether x is x. Therefore x and y are not exactly alike in every respect. The latter lacks x's feature that there is a fact concerning whether x is it. Therefore, by Leibniz's law, x and y are distinct. In that case, there is a fact after all concerning whether x is y.

[15] Kripke, "Russell's Notion of Scope," p. 1025.

So I am. The attempted refutations fail. But their failure casts no doubt on my account of the semantic contents of λ-abstracts. The claim that propositions p and q are not the same—as I make, for example, in connection with (A) and (B)—entails nothing whatsoever about whether p and q differ in truth-value, or even in modal status, or even in logical content. It is essential to my view that λ-converts like (A) and (B) are logically equivalent but nonsynonymous. In particular, on my account '$x = x$' and '$(\lambda y)[y = y](x)$' are equivalent. (Indeed, both are logically valid.) It follows that the propositions expressed (under the assignment of a value to 'x'), although distinct, share the same metaphysical status. In particular, one is a necessary truth if and only if the other is. Insofar as the objector to the necessity of identity wishes to conform to my account of abstraction, the concession that the self-identity of x is a necessary truth about x is self-defeating. Likewise, the concession that there is a fact of the matter concerning whether x is self-identical is one concession too many. The objector to the proof of the determinacy of identity is thereby committed, on my account, to the equivalent observation that there is equally a fact of the matter concerning whether x is x. The attempted refutations collapse; the proofs go through without a hitch.

Of course, a different objector to the necessity of identity might uphold part of my account of the content of abstraction while disrespecting the rest by holding that the identity of x with x is neither necessary nor equivalent to the self-identity of x (which is agreed to be necessary). Such a stance is transparently unacceptable. But this is no objection to my account, which explicitly rejects the stance in question. My account entails that, in light of their equivalence, the identity of x with x and the self-identity of x cannot differ in modal status: both are necessary. This account offers no comfort or solace to the envisioned enemy.

Indeed, the proof of the modal theorem is also (among other things) a disproof of the offending stance. And the proof is perfectly compatible with my account of the content of abstraction. The proofs of both the necessity and the determinacy of identity in fact make no assumptions whatsoever concerning the content of abstraction. They do not even invoke abstraction.

A stickler about Leibniz's law might insist otherwise. It is not implausible that, insofar as Leibniz's law is based upon Leibniz's observation that things that are one and the same are exactly alike in every respect (the indiscernibility of identicals), the substitution rule should be restricted to monadic predications. If the λ-abstraction operator is available, such a restriction is not as severe as it might seem. To guarantee validity of the proofs under such a restriction, both λ-expansion and λ-contraction would be required. From '$\Box(x = x)$', one must first prove '$(\lambda z)[\Box(x = z)](x)$' by λ-expansion before the restricted version of Leibniz's law can be applied. One will also need to perform λ-contraction on '$(\lambda z)[\Box(x = z)](y)$' to obtain '$\Box(x = y)$'. Exactly analogous modifications might also be required in the proof of the determinacy of identity. But all these modifications are easily accommodated if one is going to be that way about it.[16] The legitimacy of the proofs requires only that these instances of λ-conversion preserve truth in any model.[17] There simply is no further requirement that they should also preserve semantic content. Indeed, such a further requirement in general would be crippling to logic.

II

A second objection of Kripke's is that, in drawing a distinction in content between (A) and (B), I am committed to an implausible proliferation of propositions. For if my account is correct, Kripke argues, each of the following sentences expresses a proposition closely related to but distinct from that expressed by each of the others:

$$
\begin{array}{ll}
 & \phi_a \\
\Pi_1(a): & (\lambda x_1)[\phi_{x1}](a) \\
\Pi_2(a): & (\lambda x_2)[(\lambda x_1)[\phi_{x1}](x_2)](a) \\
\Pi_3(a): & (\lambda x_3)[(\lambda x_2)[(\lambda x_1)[\phi_{x1}](x_2)](x_3)](a) \\
 & \cdot \\
 & \cdot \\
\Pi_{i+1}(a): & (\lambda x_{i+1})[\Pi_i(x_{i+1})](a) \\
 & \cdot
\end{array}
$$

Kripke asks, "Is all this really plausible?"[18]

[16] Details are provided in my "Identity Facts," at pp. 170–71, 176 of *Metaphysics, Mathematics, and Meaning*.

[17] Some modal instances of unrestricted λ-conversion are invalid. It is illegitimate, for example, to infer by λ-expansion from '□(the number of planets is exactly how many planets there are)' to '$(\lambda n)[\Box(n$ is exactly how many planets there are)](the number of planets)' (that is, 'The number of planets is such that, necessarily, *that many* is exactly how many planets there are'). It is equally illegitimate to infer by λ-contraction from '$(\lambda n)[\Diamond\sim(n$ is exactly how many planets there are)](the number of planets)' to '◊~(the number of planets is exactly how many planets there are)'. Leibniz's law must be restricted to disallow substitution between 'the number of planets' and 'eight' in modal contexts. (Or is it 'nine'? Modal logic does not say.) By contrast, the instances of λ-conversion involved in the proofs of both the necessity and the determinacy of identity involve only abstraction on a variable (a rigid designator) and are consequently valid.

[18] Kripke, "Russell's Notion of Scope," p. 1025. 'Π_1' abbreviates the λ-abstracted predicate $\ulcorner(\lambda x_1)[\phi_{x1}]\urcorner$. For $i > 1$, 'Π_i' abbreviates the complex predicate $\ulcorner(\lambda x_i)[\ldots(\lambda x_3)[(\lambda x_2)[(\lambda x_1)[\phi_{x1}](x_2)](x_3)]\ldots(x_i)]\urcorner$. (Kripke does not use abbreviations for the λ-abstracted predicates.)

The short answer is that it is not especially implausible. (Perhaps the question is rhetorical.) A more significant question is this: Is it really so clear that the sentences in Kripke's sequence all express exactly the same proposition?

Kripke does not provide the rationale for his apparent inclination to judge that the sentences do express exactly the same proposition. The pattern of progression would seem to suggest that either all of the sentences express the same proposition or none of them do. If it is denied that the initial sentence, ϕ_a, expresses the same proposition as its successor, $\ulcorner \Pi_1(a) \urcorner$, then by parity of form, Kripke might have reasoned, it must also be denied that $\ulcorner \Pi_1(a) \urcorner$ and its own successor, $\ulcorner \Pi_2(a) \urcorner$, express the same proposition, and also $\ulcorner \Pi_2(a) \urcorner$ and its successor, $\ulcorner \Pi_3(a) \urcorner$, and so on. The opposite judgment that each consecutive pair of sentences, $\ulcorner \Pi_i(a) \urcorner$ and $\ulcorner \Pi_{i+1}(a) \urcorner$, express the same proposition is encouraged by the fact that their logical equivalence is utterly trivial. Even more important, in addition to sharing a common logical content, all of the sentences $\ulcorner \Pi_i(a) \urcorner$, for $i \geq 1$, also share a common logical *form*, consisting of a compound monadic predicate attached to the particular name '*a*'. Furthermore, each compound predicate Π_{i+1} is trivially logically equivalent to its predecessor Π_i, and thus they evidently express the very same property. Perhaps Kripke infers from these commonalities by mathematical induction that *all* of the sentences express the very proposition expressed by the initial sentence ϕ_a.

The elements of Kripke's sequence are all logically equivalent, indeed λ-convertible. That is not in dispute. I explicitly distinguish particular instances of the initial pair, ϕ_a and $\ulcorner \Pi_1(a) \urcorner$, as expressing different propositions despite being λ-convertible. It is precisely this to which Kripke objects. It does not immediately follow from my making this distinction, however, that I am committed to distinguishing likewise between $\ulcorner \Pi_1(a) \urcorner$ and $\ulcorner \Pi_2(a) \urcorner$ in regard to semantic content, nor indeed between any pairing of $\ulcorner \Pi_i(a) \urcorner$ and $\ulcorner \Pi_j(a) \urcorner$ for $i, j \geq 1$. The initial sentence, ϕ_a, unlike the rest of the sequence, need not in general have the special logical form of a predication $\ulcorner \Pi_0(a) \urcorner$ with Π_0 a monadic predicate. The asymmetry between the initial pair—which I explicitly claim need not be synonymous—and the other consecutive pairs, $\ulcorner \Pi_i(a) \urcorner$ and $\ulcorner \Pi_{i+1}(a) \urcorner$ (for $i \geq 1$), leaves ample room for a distinction between the two sorts of cases as regards the issue of preservation of semantic content.

One important basis for distinguishing the semantic contents of the initial pair is precisely the fact that $\ulcorner \Pi_1(a) \urcorner$ is a monadic predication, whereas ϕ_a need not be. This justification is utterly lacking with all the other pairs. The supposed parity of form is an illusion generated by a hasty overgeneralization. Any inclination one might have to see the relation between ϕ_a and $\ulcorner\Pi_1(a)\urcorner$ as fully on a par with that between $\ulcorner\Pi_1(a)\urcorner$ and $\ulcorner\Pi_2(a)\urcorner$ might be traced to a tempting, albeit unjustified, focus on the very special case where ϕ_a is $\ulcorner\Pi_0(a)\urcorner$ with Π_0 a monadic predicate. Admittedly, this special case is not one that motivates the general distinction in content between ϕ_a and $\ulcorner(\lambda x_1)[\phi_{x1}](a)\urcorner$, or at least not very forcefully. Perhaps it is arguable that the compound monadic predicate $\ulcorner(\lambda x_1)[\Pi_0(x_1)]\urcorner$ is exactly synonymous with Π_0 itself, making $\ulcorner(\lambda x_1)[\Pi_0(x_1)](a)\urcorner$ exactly synonymous with $\ulcorner\Pi_0(a)\urcorner$.[19] No such argument is forthcoming with regard to the general case of $\ulcorner\Pi_1(a)\urcorner$ and ϕ_a.[20]

There is an ironic analogy that aptly illustrates this point. In another context entirely, Kripke distinguishes between the propositions expressed by a pair of sentences, $\ulcorner\psi[(\iota x)\phi_x]\urcorner$ and $\ulcorner\psi(\alpha)\urcorner$, where these sentences are related in that α is a name whose "reference is fixed" by the definite description $\ulcorner(\iota x)\phi_x\urcorner$. Kripke concedes that the identity statement $\ulcorner\alpha = (\iota x)\phi_x\urcorner$ is then *a priori*, so that there is an epistemological equivalence between $\ulcorner\psi[(\iota x)\phi_x]\urcorner$ and $\ulcorner\psi(\alpha)\urcorner$. That is, their biconditional is *a priori*, according to Kripke, so that one of the sentences is *a priori* if and only if the other is as well.[21] He nevertheless distinguishes between the two sentences in regard to semantic content, on the ground that they can, and often do, differ in modal status despite their epistemological equivalence. One of Kripke's examples is the pair 'The planet that gravitationally perturbs the orbit of Uranus exerts an attractive force on Uranus' and 'Neptune exerts an attractive force on Uranus'. The former is true with respect to every

[19] This argument is rejected on Alternative (0). Anderson points out that in "A Revised Formulation of the Logic of Sense and Denotation. Alternative (1)," Church explicitly considers expanding the rules of λ-conversion to include substitution of $\ulcorner(\lambda\alpha)[\Pi(\alpha)]\urcorner$ by Π and *vice versa*, calling the resulting modification of Alternative (1) as a criterion for synonymy 'Alternative (1′)' (p. 149). Alternative (1′) identifies the semantic contents of the predicates 'is seaworthy' and '$(\lambda x)[x$ is seaworthy$]$' (in English, 'is seaworthy' and 'is a thing that is seaworthy').

[20] Anderson points out that there is another sequence of sentences analogous to Kripke's in which each consecutive pair are trivially logically equivalent: $\ulcorner\Pi_0(a)\urcorner$; $\ulcorner(\lambda F)[Fa](\Pi_0)\urcorner$; $\ulcorner(\lambda\Phi)[\Phi(\Pi_0)]((\lambda F)[Fa])\urcorner$; and so on. Here it is clear that the sentences do not all express a single proposition. Each successive sentence is constructed from components of higher type than the corresponding components of its predecessor.

[21] This aspect of Kripke's account has been disputed. Cf. my "How to Measure the Standard Metre," *Proceedings of the Aristotelian Society*, n.s., LXXXVIII (1987–1988): 193–217; reprinted in my *Content, Cognition, and Communication*, pp. 139–56. Kripke has modified his view somewhat in light of the criticism. See his "A Puzzle about Belief," in Avishai Margalit, ed., *Meaning and Use* (Boston: D. Reidel, 1979), pp. 239–83, especially at p. 281n44; reprinted in Salmon and Soames, eds., *op. cit.*, pp. 102–48, at p. 147n44; also reprinted in Matthew Davidson, ed., *On Sense and Direct Reference* (Boston: McGraw-Hill, 2007), pp. 1002–36, at p. 1034n44.

possible world in which the orbit of Uranus is perturbed by a planet. The latter is not.[22]

Imagine now a clever critic who raises the following objection against Kripke:

> Let us introduce an infinite sequence of names: α_1, α_2, α_3, ... (for example, 'Neptune$_1$', 'Neptune$_2$', and so on). We let the reference of α_1 be fixed by the description $\ulcorner(\iota x)\phi_x\urcorner$, and for each $i \geq 1$ we let the reference of α_{i+1} be fixed by its predecessor α_i. We now construct a corresponding infinite sequence of sentences: $\ulcorner\psi[(\iota x)\phi_x]\urcorner$, $\ulcorner\psi(\alpha_1)\urcorner$, $\ulcorner\psi(\alpha_2)\urcorner$, $\ulcorner\psi(\alpha_3)\urcorner$, ... Kripke concedes that these sentences are all epistemologically equivalent. If he is right, each of these sentences nevertheless expresses a proposition that is distinct from that expressed by each of the others. This is quite implausible. Rather, each of the sentences evidently expresses the very same proposition. Kripke's semantic distinction between the initial pair, $\ulcorner\psi[(\iota x)\phi_x]\urcorner$ and $\ulcorner\psi(\alpha_1)\urcorner$, is therefore bogus, a distinction without a difference.

The objection is certainly misguided. In distinguishing between the initial pair as regards the propositions expressed, Kripke is in no way committed to claiming that each of these sentences expresses a unique proposition distinct from those expressed by the others. Quite the contrary, there is a glaring difference between the specially introduced name that Kripke proposes—the special case of α_1—and all the succeeding names α_{i+1} proposed by the critic. The reference of the initial name α_1 is fixed by a *description*, $\ulcorner(\iota x)\phi_x\urcorner$—which is typically nonrigid—whereas the reference of any name α_i with $i > 1$ is fixed by *another proper name just like it.* The description that fixes the reference of α_1 is not just another *proper name*, α_0. Such a name would be rigid even when the description is not. This is precisely the point. The ground for distinguishing the initial pair of sentences—their potential for differing in modal status—is consequently utterly lacking with any pairing of the $\ulcorner\psi(\alpha_i)\urcorner$.

The situation here is exactly analogous to Kripke's own objection to my account of the content of abstraction, but for the substitution of modal for structural properties.

III

I have argued against Kripke that my distinguishing ϕ_a and $\ulcorner\Pi_1(a)\urcorner$ as regards content does not commit me to distinguishing the rest. But the mere fact that there is conceptual room for a particular position does not entail that the position is sound.

[22] Cf. Kripke's *Naming and Necessity* (Cambridge: Harvard, 1972, 1980), at pp. 14–15, 79n.

Let us return for a moment to the special case where ϕ_a is $\ulcorner\Pi_0(a)\urcorner$ with Π_0 a monadic predicate. Is the contention correct that $\ulcorner(\lambda x_1)[\Pi_0(x_1)](a)\urcorner$ expresses exactly the same proposition as $\ulcorner\Pi_0(a)\urcorner$? Consider an English analog of this. Let Π_0 be the predicate 'is seaworthy', and consider the expanded sentence '*a* is a thing that is seaworthy'. Is this strictly synonymous with the contracted '*a* is seaworthy'? An exactly similar question arises about the semantic contents of $\ulcorner\Pi_2(a)\urcorner$, $\ulcorner\Pi_3(a)\urcorner$, $\ulcorner\Pi_4(a)\urcorner$, and so on. Is the proposition that *a* is a thing that is a thing that is seaworthy (the content of $\ulcorner\Pi_2(a)\urcorner$) exactly the same as the proposition that *a* is seaworthy? What about the proposition that *a* is a thing that is a thing that is a thing that is seaworthy ($\ulcorner\Pi_3(a)\urcorner$)? Are these all exactly the same, as Kripke is inclined to believe?

These are not at all easy questions with easy answers. Kripke cannot legitimately claim that intuition squarely supports his favored judgment that they are all a single proposition. This verdict is no more intuitive than its rival. As already noted, all these sentences do have a common logical content and a common logical form, consisting of a monadic predicate attached to '*a*'. And the predicates are all trivially equivalent. For this reason, I am prepared to assert that they stand or fall together. Either no two of them express exactly the same proposition, or they all do; it is either all or none. On the other hand, considerations of propositional attitude should give one serious pause before declaring that these sentences express exactly the same proposition with not a hair's difference, as Kripke implicitly does. It should be noted also that on Alternative (0) these same sentences are deemed not strictly synonymous.[23] I have taken no explicit stand on the issue. I am somewhat inclined to judge that each does indeed express a unique proposition distinct from the others, but I am prepared to be persuaded either way.

Even if Kripke's implicitly proposed identification in content between ϕ_a and $\ulcorner\Pi_1(a)\urcorner$ is correct in the special case where ϕ_a is $\ulcorner\Pi_0(a)\urcorner$ for some monadic predicate Π_0—a big 'if'—it cannot be argued plausibly that such identification extends to the case where ϕ_a does not have this special form. We have already seen this in one particular case, where ϕ_a is (A) and $\ulcorner\Pi_1(a)\urcorner$ is (B).[24]

[23] See note 19 above. Church says of the proposed modification of Alternative (1) mentioned there—allowing that interchange between $\ulcorner(\lambda\alpha)[\Pi(\alpha)]\urcorner$ and Π preserves sense—that whether such modification is needed "in the end…may depend on somewhat doubtful judgments as to whether given (declarative) sentences convey exactly the same item of information, or whether instead it is closely related but different such items" (pp. 148–49). Interestingly, Church does not consider emending Alternative (0) similarly.

[24] Does Kripke simply overlook the case in which ϕ_a involves multiple occurrences of '*a*'? He explicitly points out, "If *n*-place relations are involved, the situation comes

Consider also the case where ϕ_a is a simple conjunction, for example, '*a* is large and *a* is seaworthy'. Here, $\ulcorner \Pi_1(a) \urcorner$ expresses the monadic-predication proposition that *a* is a thing that is both large and seaworthy. The next sentence, $\ulcorner \Pi_2(a) \urcorner$, expresses that *a* is a thing that is a thing that is both large and seaworthy, the next that *a* is a thing that is a thing that is a thing that is both large and seaworthy, and so on. Are all these propositions, strictly speaking, different propositions from one another? Are they all one and the very same proposition? I do not endorse either verdict. What I do claim is that, whatever Kripke's inclinations might be, the conjunctive proposition that *a* is large and also *a* is seaworthy is surely different from (even though logically equivalent to) the monadic-predication proposition that *a* is a thing that is both-large-and-seaworthy. One who is mistaken about the identity of the yacht in question might well believe the former proposition ("This yacht is large whereas that yacht is seaworthy") without thereby believing the latter.[25]

I also claim that '$(\lambda x)[x$ is large & x is seaworthy$](a)$' expresses the latter proposition rather than the former. Similarly with regard to the contents of (A_E) and (B_E). And I claim that intuition decidedly supports these judgments. Even if these verdicts, or their grounds, commit me to distinguishing each of the sentences in the sequence as regards semantic content, as Kripke asserts, the latter issue is far too delicate to decide the significantly easier issue concerning ϕ_a and $\ulcorner \Pi_1(a) \urcorner$ when ϕ_a involves significant structure (for example, multiple occurrences of '*a*')—an issue that is decided on philosophically intuitive grounds through consideration of a range of such cases.

IV

Kripke's most forceful counter-consideration concerns Russell's account of (B) in contrast to mine. Having once embraced the classical

---

to involve complicated infinite trees." It is precisely in such cases that the potential non-synonymy of ϕ_a and $\ulcorner \Pi_1(a) \urcorner$ is laid bare.

Kripke's observation is strictly true even for the special case where ϕ_a is of the form $\ulcorner \Pi_0(a) \urcorner$ with Π_0 a monadic predicate. This sentence yields not only $\ulcorner(\lambda x)[\Pi_0(x)](a)\urcorner$ by λ-expansion, but also $\ulcorner(\lambda x)[\Pi_0(a)](a)\urcorner$. Is the proposition that *a* is a thing that is seaworthy the same as, or different from, the proposition that *a* is such that *a* is seaworthy? Is either the same as the (logically equivalent) proposition that *a* is seaworthy? Do considerations of propositional attitude shed any light?

[25] Distinguishing the contents of '*a* is large & *a* is seaworthy' and its λ-convert, '$(\lambda x)[x$ is large & x is seaworthy$](a)$', provides the beginning of an explanation of why it is that one who believes the content of the former might not yet be in a position to be able to infer that *a* is a thing both large and seaworthy. The situation here is exactly like Kripke's own case of Pierre, in "A Puzzle about Belief," *loc. cit.* See notes 8 and 21 above.

Fregean distinction between semantic content ("meaning") and designation, by 1905 Russell came to reject it.[26] He regarded sentences, roughly, as designators of propositions. Furthermore, he regarded λ-abstracts like '$(\lambda x)[x$ is larger than x is$]$' as functors that designate propositional functions. His view supports a stronger version of λ-conversion rules, on which a sentence and its λ-convert are not only alike in truth-value but also co-designative on logical grounds alone. In particular, on Russell's view, (B) designates the proposition obtained by applying the function designated by '$(\lambda x)[x$ is larger than x is$]$' to the yacht designated by 'a'. This, on Russell's account, is exactly the proposition designated by (A).[27] Thus, as Kripke puts it, Russell did not intend any distinction between (A) and (B).

Kripke says that Russell's account of (A) and (B) strikes him as correct. He adds, "Nor does a mathematician analogously intend any distinction between $\lambda x(x!)(3)$ and the number 6. Nor did Church, inventor of the lambda notation, intend any such distinction."[28] It must be noted in response that, quite the contrary, the mathematician's understanding and use of the λ-calculus in fact casts serious doubt on Russell's account of (A) and (B) while simultaneously supporting my account. The result of applying the factorial function $(\lambda x)[x!]$ to the number three is indeed the number six. The 'is' here is the 'is' of strict identity. The complex expression '$(\lambda x)[x!](3)$' and the numeral '6' are co-designative, hence co-extensional. Indeed, they are mathematically equivalent, in the sense that '$(\lambda x)[x!](3) = 6$' is a mathematical theorem. But as Church would have observed, these two expressions for the number six differ dramatically in semantic content, in "sense." To begin with, although they are mathematically equivalent, the two expressions are arguably not logically equivalent, let alone λ-convertible. If so, they are not deemed synonymous even on Alternative (2), let alone on Alternatives (1) or (0). This is a striking point of disanalogy with (A) and (B), which are λ-converts.[29]

[26] While calling the distinction 'Fregean', it should be acknowledged that philosophers before Frege also upheld the distinction—including Russell's godfather, John Stuart Mill, who distinguished "connotation" from "denotation." Russell himself also upheld the distinction, before his 1905 breakthrough.

[27] See note 2 above. Russell's account is obtained from the explanation provided there by assuming that formulae designate propositions, under assignments of values to variables, and that where ϕ_α is a formula, the designatum, with respect to an assignment s of values to variables, of the λ-abstract $\ulcorner(\lambda\alpha)[\phi_\alpha]\urcorner$ is the function that assigns to anything o in the range of the variable α, the designatum of ϕ_α with respect to the assignment that assigns o to α and is otherwise the same as s.

[28] Kripke, "Russell's Notion of Scope," p. 1025.

[29] I assume that the numeral '6' is an individual constant, defined perhaps by 'the successor of 5' or taken as primitive. An example better suited to Kripke's purpose

Even aside from such considerations, the nonsynonymy of '$(\lambda x)[x!]$ (3)' and '6' is obvious. The numeral '6' is the canonical designator of six in Arabic notation. One does not understand the numeral—one does not grasp what its semantic content is—unless one knows which number is designated thereby. By contrast, calculation is required to determine that '$\lambda x[x!](3)$' designates six. These mathematically equivalent designators are thus co-extensional but nonsynonymous. On my account, (A) and (B) are analogously logically equivalent, hence co-extensional, yet nonsynonymous. This is not so on Russell's account, which wrongly deems (A) and (B) strictly synonymous. I embrace Kripke's analogy, modulo that different species of equivalence are involved. Contrary to the thrust of Kripke's objections, however, the defender of Alternative (1) cannot do so.

It might be replied that, just as classical logic is concerned with extensions and not with intensions, so pure mathematics is concerned only with the extensions of mathematical notation. The mathematician *qua* mathematician writes the equation '$3! = 6$' but does not also say whether this sentence is informative, or trivial, or according to Kant to be labeled 'analytic', or containing a very valuable extension of knowledge. When Frege addressed these matters, he was wearing his philosopher hat, not his mathematician hat. In particular, pure mathematics does not say whether '$3! = 6$' differs in semantic content, or in cognitive value, from '$6 = 6$'. Strictly speaking, it may be argued, pure mathematics does not say anything at all about particular expressions ('3!', '6', '$\lambda x[x!](3)$', and so on). In particular, the mathematician *qua* mathematician does not draw any semantic distinction between '$\lambda x[x!](3)$' and '6'.

I do not believe this reply accords with Kripke's intent, but it requires rebuttal all the same. Gödel's method of arithmetizing syntax already demonstrates that there is a sense in which pure mathematics is not altogether free of reference to the very notation in which its theorems are formulated. Some mathematical theorems explicitly concern overtly semantic notions—for example, Tarski's theorem about truth. The very point of Church's *LSD* is that it is a mathematical theory about the entities that serve as semantic contents. Church's intention was undoubtedly that this mathematical theory ultimately was to be combined with a mathematical theory of structures (models) to yield a mathematical theory that does for the

---

would be something like '$(\lambda x)[2(x!) - x^2 + x](3)$', as contrasted with its λ-convert '$2(3!) - 3^2 + 3$'—except that such genuinely analogous examples strongly support my account of (A) and (B) over the Russellian account that Kripke favors. (See note 37 below.)

semantics of sense (content) roughly what Tarski's theory of truth-in-a-model does for extensional semantics.

Furthermore, it is simply false that the mathematician (*qua* mathematician) does not discriminate among different mathematical expressions for the number six. It is a mathematical question, for example, how to designate six using a sequence of occurrences of '1' and '0' ("ones and zeroes") in binary notation. (It is not an interesting mathematical question, but it is a mathematical question.) There is no similar mathematical issue of how to designate six in Arabic notation. (There is an issue, but it is not a mathematical one.)

More to the point, when the mathematician (*qua* mathematician) endeavors to calculate the factorial function for the number three as argument—a purely mathematical task if anything is—there are infinitely many expressions for six that do *not* qualify as solutions. One expression that is absolutely disqualified is '$(\lambda x)[x!](3)$'. In fact, among expressions for six, this one, along with its λ-convert '3!', are among the least qualified of all. (Other expressions for six that are also absolutely disqualified include '$3! + 0$', '$3! - 0$', '$3!/1$', and so on.) The most qualified expression in Arabic notation—maybe the only absolutely and fully qualified expression—is precisely the Arabic numeral '6'. Thus, although the mathematician draws no distinction between six and the result of applying the factorial function to three, the mathematician (*qua* mathematician) draws a very sharp distinction between '6' and '$(\lambda x)[x!](3)$'.[30] Among expressions for six, those that qualify as solutions to the equation '$x = 3!$' and those that do not are separate but equal. By definition, this discrimination—the bisection of expressions for six into disjoint sets of those that qualify as solutions to the factorial of three and those that do not—is not based solely on designation. It is based, in fact, on semantic content. The mathematician *qua* mathematician, whether he/she realizes it or not, is thereby concerned with matters of semantic content, even if only implicitly.

Though the mathematician might not explicitly mention the semantic contents of mathematical expressions, it is plainly a *fact* (even if it is not asserted) that the mathematically equivalent expressions '$(\lambda x)[x!](3)$' and '6' differ in semantic content. They are mathematically equivalent and therefore co-extensional, but they are not strictly synonymous (nor λ-convertible).

[30] In his 1992 Alfred North Whitehead Lectures on "Logicism, Wittgenstein, and *De Re* Beliefs about Numbers" at Harvard University, Kripke made a closely related point, specifically concerning recursive-function theory.

V

Kripke's depiction of Church is historically inaccurate. Church explicitly preferred Alternative (0) to the less strict Alternatives (1) and (2) as an explication of the objects of the propositional attitudes. He wrote that he "attaches the greater importance to Alternative (0) [than to Alternative (1)] because it would seem that it is in this direction... that a satisfactory analysis is to be sought of statements regarding assertion and belief."[31] This characteristically judicious comparison of the relative merits of the alternatives is exactly correct. Whether Alternative (0) is a step in the right direction or not, certainly Alternatives (1) and (2) go in the wrong direction.

Further, in correspondence with C. Anthony Anderson in 1973, Church argued that Alternative (1) is unsuitable for a "logic of belief statements." Church notes there that the apparatus in his paper, "A Formulation of the Simple Theory of Types," provides a notation suitable for arithmetic—with canonical designators for the natural numbers, a sign for multiplication, and so on—in which the notation, **k×l**, for the result of multiplying the numbers canonically designated by **k** and **l**, is in general λ-convertible with the canonical designator, **m**, of the resulting product.[32] Nevertheless, Church argues, someone might erroneously believe the proposition expressed by ⌜**m** is prime⌝ without thereby believing the proposition expressed by ⌜**k×l** is prime⌝.[33] (See also note 23 above.) Some twenty years later, Church said that taking λ-convertibility as a criterion for synonymy "may be thought counterintuitive if propositions in the sense of Alternative (1) are to be taken as objects of assertion and belief."[34]

Although they are logically equivalent and, indeed, λ-convertible, (A) and (B) are not *strictly* synonymous on Alternative (0). Church regarded (A) and (B) as expressing different propositions. Curiously, Kripke's example and his appeal to Church's authority thus strongly support my account of the semantics of (A) and (B) over the Russellian account that Kripke prefers, while these same considerations in fact discredit the latter.[35]

[31] Church, "A Formulation of *LSD*," at p. 7n7.

[32] Church, "A Formulation of the Simple Theory of Types," *Journal of Symbolic Logic*, v, 2 (June 1940): 56–68.

[33] Cf. Anderson, "Alternative (1*): A Criterion of Identity for Intensional Entities," at pp. 421–23. I am grateful to Anderson for sharing with me Church's letter, dated May 5, 1973. I have taken liberties in adapting Church's formulation to the notation of the present essay. The canonical number-designators in the Church apparatus are not in general synonymous with their Arabic-numeral counterparts. See note 29 above.

[34] Church, "A Revised Formulation of the Logic of Sense and Denotation. Alternative (1)," at p. 156n2.

[35] Whereas Kripke evidently misunderstands Church's position, he might well have Russell correct. If so, then Russell evidently changed his mind after *The Principles of*

The argument for the nonsynonymy of (A) and (B) by analogy to '$(\lambda x)[x!](3)$' and '6' exploits the Fregean semantic distinction between content (*Sinn*) and extension (*Bedeutung*). In the short passage quoted in the preceding section, Kripke is pointing out that "the mathematician," and in particular Church, holds that the factorial of three just *is* six, so that '$(\lambda x)[x!](3)$' and '6' are co-designative. Kripke is not claiming that mathematicians regard these expressions as synonyms. By analogy, Kripke believes (A) and (B) should be seen as designating the same proposition.

Church indeed did hold, with Frege, that (A) and (B) may be seen as co-designative—but of a truth-value, not of a proposition. In this, I fully agree with Frege and Church and disagree with Russell and Kripke. If sentences are designators, they designate truth-values. Propositions are the semantic contents of sentences, not the designata. Russell, of course, was dubious of the Fregean distinction between semantic content and designation, and strove to avoid it. As is well known, Kripke too finds much in Frege's distinction to dispute, especially in regard to proper names. While I am completely convinced by Kripke's critique of the Fregean distinction as applied to proper names, I am equally convinced that it is a mistake to follow Russell rather than Frege-Church with regard to sentences.

Let us shift attention for a moment from sentences to singular terms. The mathematical analog of a Russellian propositional function in connection with the factorial function is a function that assigns to a

---

Mathematics (Cambridge: University Press, 1903). An adherent of Alternative (1) will hold that '$(\lambda xy)[x > y](ab)$' and '$(\lambda yx)[x > y](ba)$' are synonymous, both expressing simply that *a* is greater than *b*. Anderson points out, however, that in *The Principles* Russell takes up the question, deciding that these are not synonymous. Russell writes:

> A question of considerable importance to logic, and especially to the theory of inference, may be raised with regard to difference of sense. Are $a\boldsymbol{R}b$ and $b\check{\boldsymbol{R}}a$ really different propositions, or do they only differ linguistically? It may be held that there is only one relation $\boldsymbol{R}$, and that all necessary distinctions can be obtained from that between $a\boldsymbol{R}b$ and $b\boldsymbol{R}a$. It may be said that, owing to the exigencies of speech and writing, we are compelled to mention either *a* or *b* first, and that this gives a seeming difference between "*a* is greater than *b*" and "*b* is less than *a*"; but that, in reality, these two propositions are identical. But if we take this view we shall find it hard to explain the indubitable distinction between *greater* and *less*. These two words have certainly each a meaning, even when no terms are mentioned as related by them. And they certainly have different meanings, and are [i.e., denote] certainly relations. Hence if we are to hold that "*a* is greater than *b*" and "*b* is less than *a*" are the same proposition, we shall have to maintain that both *greater* and *less* enter into each of these propositions, which seems obviously false.... Hence, it would seem, we must admit that $\boldsymbol{R}$ and $\check{\boldsymbol{R}}$ are distinct relations...and "$a\boldsymbol{R}b$ implies $b\check{\boldsymbol{R}}a$" must be a genuine inference. (Section 219, pp. 228–29).

natural number n a particular numerical concept: *the product of the natural numbers less than or equal to* n. This function assigns to three not six, but a concept: *the product of the natural numbers less than or equal to three*. This function is decidedly not what the mathematician takes '$(\lambda x)[x!]$' to designate. The λ-abstract designates the factorial function—a mathematical function from numbers to numbers—not a function from numbers to numerical concepts involving those numbers. Insofar as the latter function is a semantic value of '$(\lambda x)[x!]$', it is a semantic value at the level of semantic content, not one at the level of designation.

Shift back to sentences. Before his immersion in *LSD*, Church had proffered a proof of sorts, from a principle of *compositionality of designation*, for the Fregean conclusion that, if sentences are designators (and if trivially mathematically equivalent designators are co-designative), then sentences that are alike in truth-value are co-designative even if they express different propositions. At nearly the same time as Church, Gödel provided a similar proof, also inspired by Frege's arguments.[36] The general form of the argument has been called *the slingshot* because of its supposedly disarming power to slay philosophical giants (for example, the thesis that sentences designate propositions). The principle of compositionality of designation is that (in the absence of such deviant devices as quotation, intensional operators, 'believes that', and so on) the designatum of a compound designator is a function of the designata of any component designators. As Anderson points out, a severely restricted version of compositionality of designation suffices for the proof:

(*Comp*) Assuming that sentences are designators, the designatum of an identity sentence $\ulcorner\alpha = \beta\urcorner$ is a function of the designata of its singular terms, α and β.

A simple formulation of the proof goes as follows: Assume (*Comp*) and that trivially mathematically equivalent designators are co-designative. It follows that, if sentences are designators, then all sentences that are alike in truth-value are co-designative. *Proof*: Assume that sentences are designators. Let ϕ and ψ be any sentences that are alike in truth-value, for example, 'Snow is white' and 'Water runs downhill' or 'Snow is green' and 'Water runs uphill'. Then ϕ and the complex sentence

[36] Church, "Review of Carnap's *Introduction to Semantics*," *The Philosophical Review*, LII, 3 (May 1943): 298–304; and *Introduction to Mathematical Logic I* (Princeton: University Press, 1956), pp. 24–25, reprinted in Davidson, ed., *op. cit.*, pp. 54–81, at p. 75; Gödel, "Russell's Mathematical Logic," in Paul Arthur Schilpp, ed., *The Philosophy of Bertrand Russell* (La Salle, IL: Open Court, 1989), pp. 123–53, at pp. 128–29.

$\ulcorner(\iota n)([\phi \supset n = 1] \wedge [\sim\phi \supset n = 0]) = 1\urcorner$ are co-designative, since they are trivially mathematically equivalent. By (*Comp*), the latter co-designates with $\ulcorner(\iota n)([\psi \supset n = 1] \wedge [\sim\psi \supset n = 0]) = 1\urcorner$, since the definite descriptions contained within the two sentences are co-designative. Since this last sentence is trivially mathematically equivalent to ψ, they are co-designative. Therefore, ϕ and ψ are co-designative.[37]

A straightforward variant of the proof shows that if predicates are designators, then co-extensional predicates are co-designative even if they are not synonymous, as with Quine's example of 'is a cordate' and 'is a renate'.[38] This result supports Church's contention that the λ-abstract, $\ulcorner(\lambda x)[\phi_x]\urcorner$, designates not a propositional function but something fully extensional—the class of things satisfying ϕ_x or alternatively, with Church (following Frege), this class's characteristic function.

Kripke knows the Church-Gödel proof well. (The particular proof presented here is derived from improved versions of the argument which David Kaplan had formulated in his 1964 doctoral dissertation on "Foundations of Intensional Logic" and which Kripke presented independently in an undergraduate course I took in 1972.) The argument's conclusion supports the thesis—which I accept—that insofar as '$(\lambda x)[x$ is larger than x is]' is a designator, it designates the empty class (more accurately, the constant function to falsehood) rather than a Russellian propositional function. To resist the Church-Gödel argument, Kripke is ultimately committed to rejecting even the minimal compositionality principle (*Comp*). This rejection strikes the present author as excessively implausible. (See the Appendix below.)

Kripke evidently believes that '$(\lambda x)[x$ is larger than x is]' designates a propositional function, which assigns to anything the proposition

[37] David Braun correctly points out that the equivalence between ϕ and $\ulcorner(\iota n)([\phi \supset n = 1] \wedge [\sim\phi \supset n = 0]) = 1\urcorner$ is not logical in the usual sense but mathematical. The bi-conditional formed from these two sentences is not a first-order-logical truth but a first-order-logical consequence of the truism, '$1 \neq 0$'. On the other hand, '$1 \neq 0$' is arguably a trivial second-order logical truth, so that the two sentences are trivially second-order-logically equivalent. Cf. my "Numbers versus Nominalists," *Analysis*, LXVIII, 3 (July 2008): 177–82.

[38] Assume that predicates are designators. Let Π and Π' be any co-extensional monadic predicates, for example, 'is a cordate' and 'is a renate'. Then Π co-designates with the complex predicate $\ulcorner(\lambda x)[(\iota n)([\Pi(x) \supset n = 1] \wedge [\sim\Pi(x) \supset n = 0]) = 1]\urcorner$, since they are trivially mathematically equivalent. By compositionality of designation, the latter predicate co-designates with $\ulcorner(\lambda x)[(\iota n)([\Pi'(x) \supset n = 1] \wedge [\sim\Pi'(x) \supset n = 0]) = 1]\urcorner$ since for every value of 'x', the definite descriptions contained within the two complex predicates are co-designative. Since this last predicate is trivially mathematically equivalent to Π', they are co-designative. Therefore, Π and Π' are co-designative. Cf. my *Reference and Essence*, pp. 48–52; and my *Frege's Puzzle* (Atascadero, CA: Ridgeview, 1986, 1991), pp. 22–23.

that the thing in question is larger than that same thing is, rather than the proposition that it is self-larger. In light of the Church-Gödel argument, it is rather the semantic content of the λ-abstract, not the designatum, which is closely related to a Russellian propositional function. I do not deny that the propositional function is a semantic value of the λ-abstract. On the contrary, I would insist that it obviously *is* a semantic value of some sort—at the level of semantic content, not at the level of extension. What I deny is that the semantic content is just this function. It is significantly more plausible that (B) expresses the monadic proposition that *a* is self-larger rather than the binary-relational proposition that *a* is larger than *a* is. Kripke's misgivings notwithstanding, I know of no convincing reason to doubt this. And there are good reasons to acknowledge it.

On the other side, advocacy of Alternative (1), and rejection of (*Comp*), seems to be based on confusion. It is possible using the λ-operator to designate a singular proposition as the value of a propositional function. The following expression presents the proposition that *a* is larger than *a* is precisely as the result of applying to *a* the very propositional function mentioned in the preceding paragraph:

(C) (λx)[the proposition that x is larger than x is](a).

(See note 2.) This expression thus designates the very proposition expressed by (A). But (C) is not thereby synonymous with (A). Let alone is (C) synonymous with (B), which expresses a different proposition from the one that (A) expresses and (C) designates. In fact, (C) does not *express* a proposition at all. It is not a sentence; it is a term. It designates the proposition expressed by (A) by *describing* it, as the value of a particular propositional function for a specified argument. As such, (C) is a compound descriptive term for a proposition, in the same way that '$(\lambda x)[x!](3)$' is a compound descriptive term for six. The content of (C) is not a proposition; it is a proposition concept. This contrasts rather sharply with (B).

The λ-abstraction operator is essentially a device for forming compound functors and other operators from open expressions, in particular, compound predicates from open formulae. A predicate, whether simple or compound, expresses an attribute or concept as its semantic content, one that determines a class as extension.[39] The predicate's content is a component of the contents of the typical sentences in which the predicate occurs. A predication sentence, $\ulcorner\Pi(\alpha_1, \alpha_2, \ldots\alpha_n)\urcorner$, thereby expresses a content of a special sort peculiar to predication sentences.

[39] This observation might be credited to John Stuart Mill, who observed that a "general name" is invariably "connotative" as well as "denotative." See note 26 above.

Were it not that its predicate expresses a class-determining attribute or concept, a predication sentence would not express a proposition. If Π were simply a Millian name (a logically proper name) for a class, for example, without connotation (intensional content), then the string of symbols, $\ulcorner\Pi(\alpha)\urcorner$, would not express a proposition; it would not *assert* anything. Alternative (1) is essentially blind to the special predication role of the λ-abstracted predicate. In its blindness, Alternative (1) reads (B) as a descriptive designation of a proposition rather than as a predication sentence. In effect, Alternative (1) mistakes (B) for a notational variant of (C)—or worse, for a proposition name whose designatum is fixed by a description, *to wit*, by (C).

NATHAN SALMON

University of California, Santa Barbara and
City University of New York Graduate Center

APPENDIX: COMPOSITIONALITY AND THE CHURCH-GÖDEL ARGUMENT

Russell's theory of descriptions endeavors to sidestep the Church-Gödel "proof" concerning the designata of sentences. That theory contradicts the argument's plausible assumption that a definite description $\ulcorner(\iota\alpha)\phi_\alpha\urcorner$ designates the thing that satisfies its matrix ϕ_α if exactly one thing does. This attempt to block the Church-Gödel argument is ultimately inadequate. The Fregean assumption that a proper definite description designates the thing that uniquely answers to it is in fact inessential. Even if definite descriptions are regarded as restricted existential quantifiers (on the model of ⌜some unique *NP*⌝) rather than as singular terms, the two definite descriptions invoked in the argument are still co-extensional, so that a restricted version of extensionality of designation would still be applicable:

(*Ext*) Assuming that sentences are designators, the designatum of a sentence of the form $\ulcorner(\iota\alpha)\phi_\alpha = \beta\urcorner$ is a function of the extensions of both $\ulcorner(\iota\alpha)\phi_\alpha\urcorner$ and β.

Strictly speaking, Russell regards definite descriptions neither as singular terms nor as restricted quantifiers but as altogether content-less (having no "meaning in isolation"). I doubt Kripke wishes to go so far. (He has surprised me on occasion.) Indeed, in *Naming and Necessity*, Kripke explicitly acknowledges that a natural-language definite description is a *designator* (typically nonrigid) of the thing that uniquely answers to it.[40]

[40] Kripke is somewhat more circumspect in "Speaker's Reference and Semantic Reference," in P. French, T. Uehling, and H. Wettstein, eds., *Contemporary Perspectives in the Philosophy of Language* (Minneapolis: Minnesota UP, 1979), pp. 6–27. But see note 42 below.

Relying on extensionality of designation in lieu of (*Comp*), definite descriptions are eliminable altogether from the proof in favor of λ-abstracted predicates. The original sentence ϕ is trivially mathematically equivalent to $\ulcorner(\lambda n)[(\phi \supset n = 1) \wedge (\sim\phi \supset n = 0)](1)\urcorner$ (that is, ⌜One is a thing that is one if ϕ and is zero if not-ϕ⌝). Replacing the predicate $\ulcorner(\lambda n)[(\phi \supset n = 1) \wedge (\sim\phi \supset n = 0)]\urcorner$, in turn, by the co-extensional predicate $\ulcorner(\lambda n)[(\psi \supset n = 1) \wedge (\sim\psi \supset n = 0)]\urcorner$ yields $\ulcorner(\lambda n)[(\psi \supset n = 1) \wedge (\sim\psi \supset n = 0)](1)\urcorner$, which is trivially mathematically equivalent to ψ. The following minimal extensionality principle suffices:

> (*Ext′*) Assuming that sentences are designators, the designatum of a monadic-predication sentence, $\ulcorner\Pi(\alpha)\urcorner$, is a function of the designatum of its singular term α and the extension of its predicate Π.

Russell and Kripke (assuming they will concede that trivially mathematically equivalent designators are co-designative) are thus committed to rejecting (*Ext′*). Evidently, on their view, the designatum of a compound designator containing a predicate—even if it is a predication sentence, $\ulcorner\Pi(\alpha_1, \alpha_2, \ldots\alpha_n)\urcorner$—depends on the propositional function semantically associated with that predicate, rather than on the predicate's extension. Otherwise, the sentences '*a* is a renate' and '*a* is a cordate' will be co-designative (as Frege, Church, and I take them to be).

Whereas Russell denied that definite descriptions are contentful, he nevertheless regarded proper definite descriptions as simulating designation, whereby a proper definite description pseudo-designates the thing that uniquely answers to it. Russell used the term 'denotation' to cover (among other things) both the designatum of a singular term and the pseudo-designatum of a description. A variant of (*Comp*) is sufficient for the proof:

> (*Comp′*) Assuming that sentences are designators, the designatum of a sentence of the form $\ulcorner(\iota\alpha)\phi_\alpha = \beta\urcorner$ is a function of the "denotations" (that is, the designata or pseudo-designata) of $\ulcorner(\iota\alpha)\phi_\alpha\urcorner$ and β.

Russell and Kripke are committed to rejecting this restricted compositionality principle.

The Church-Gödel proof can make do instead with the original minimal compositionality principle (*Comp*) by using an artificial variable-binding operator—which might as well be inverted iota—and which, it is stipulated, forms a compound singular term, in contrast to a restricted quantifier, from an open formula. This device makes for the strongest version of the argument, since it relies on the weakest assumptions. Although Russell had proposed exactly such a device himself, by the time of the publication of "On Denoting" he needed to maintain that such a device is somehow impossible. He seems to have believed exactly this.[41] It is extremely doubtful that Kripke wishes to go so far. Indeed, in a discussion removed from the present one, Kripke postulated a natural-language analog to the very device in question, explicitly arguing that interpreting the English definite article 'the' by means of this device yields a language that might even be English.[42] In accepting the possibility of this device, Kripke thus rejects (*Comp*) or is committed to doing so—just as Russell must reject (*Comp′*).

[41] Cf. my account of Russell's notorious "Gray's *Elegy*" argument in "On Designating," *Mind*, CXIV, 456 (October 2005): 1069–133; reprinted in my *Metaphysics, Mathematics, and Meaning*, pp. 286–334.

[42] Cf. Kripke's so-called *weak Russell language*, set out in "Speaker's Reference and Semantic Reference." This possible language "takes definite descriptions to be primitive designators" (p. 16).

IT SIMPLY DOES NOT ADD UP: TROUBLE WITH OVERALL SIMILARITY*

Comparative overall similarity lies at the basis of a lot of recent metaphysics and epistemology. It is a poor foundation. Overall similarity is supposed to be an aggregate of similarities and differences in various respects. But there is no good way of combining them all.

I. SIMILARITIES AND DIFFERENCES

Similarity is relative: things are similar in one respect but different in another. And it is comparative: some things are more similar to each other, in a given respect, than are other things. This much is quite straightforward. The idea behind comparative overall similarity has been that some things might be more similar than other things—but simply so, not in any particular respect—somehow as a result of similarities and differences in *several* respects. This is not at all straightforward, because overall similarity is supposed to be some sort of aggregate. It is supposed to be the result of adding up similarities or weighing them against differences or combining them in another way.

Comparative overall similarity, I shall argue, does not meet the demands that philosophers make of it. At root, the trouble is that, in general, greater similarity in one respect will not make up for less similarity in another respect. For this reason, we will see, there can be no combining of the various similarities and differences of things into useful comparisons of overall similarity. Before going any further, though, let us stop and see what depends on this.

As a first example, take the question, "How could things have been different?" One theory has it that people and other ordinary things have counterparts in other possible worlds. That is how, for example, you could have grown up in a traveling circus: you could have done so because some counterpart of yours did in fact grow up in one. Now, counterparthood is a matter of overall similarity. Your counterpart is someone who resembles you more closely, overall, than do others in his possible world.[1]

* For discussion, I thank Johan van Benthem, Ruth Chang, Max Cresswell, Matti Eklund, Steve Kuhn, Barry Loewer, Ray Martin, Dan Moller, Georges Rey, Denis Robinson, Michael Slote, Mark Wunderlich, and everybody at Rosewind Farm.

[1] David Lewis, "Counterpart Theory and Quantified Modal Logic," this JOURNAL, LXV, 5 (March 1968): 113–26.

0022-362X/10/0709/469–490

Or, consider how things undergo change. This is tied up with overall similarity if, as many think, ordinary objects persist from one time to the next by having a series of temporal stages, and change by having stages that are dissimilar. It is thought that what conjoins the successive stages—that is, what makes them parts of one temporally extended thing—is not only their causal connectedness but also their overall similarity.[2] This explains why a pile of decayed planks, discarded in the process of preserving the ship of Theseus, is not itself the original ship, although the required causal connection is there. The pile's stages do not resemble earlier stages of the original ship as closely, overall, as do the stages of the preserved ship.[3]

There is yet more work cut out for overall similarity. Some say that a counterfactual conditional sentence is true if some possible world in which both the antecedent and consequent are true is more similar, overall, to the world of evaluation, than is any world in which the antecedent is true but the consequent is false.[4] Accounts of causation,[5] the direction of time,[6] knowledge,[7] and intentionality[8] in turn depend on counterfactuals. Verisimilitude, or comparative likeness of false theories to the truth, has been thought to be an aggregate of likenesses in respect of truth and of content.[9]

Nelson Goodman once complained of the many ways in which he thought similarities have failed philosophy. Similarity judgments, he observed, require not only the selection of relevant properties but also the weighting of their importance. Importance is a volatile matter,

[2] Lewis writes that temporal parts "are united as much by relations of causal dependence as by qualitative similarity" in *On the Plurality of Worlds* (New York: Blackwell, 1986), p. 218. According to Robert Nozick, temporal identity is a matter of "not merely the degree of causal connection, but also the qualitative connection of what is connected, as this is judged by some weighting of dimensions and features in a similarity metric." See his *Philosophical Explanations* (Cambridge: Harvard, 1981), p. 37.

[3] The ancient Athenians are said to have preserved the ship in which Theseus returned by taking away old planks as they decayed and replacing them with new ones. I assume that the pile of discarded planks is a continuer that, but for the presence of the preserved ship, might itself have been the original ship of Theseus.

[4] See Lewis, *Counterfactuals* (Cambridge: Harvard, 1973); and Robert Stalnaker, "A Theory of Conditionals," in Nicholas Rescher, ed., *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, 2 (Oxford: Blackwell, 1968), pp. 98–112.

[5] Lewis, "Causation," this JOURNAL, LXX, 17 (Oct. 11, 1973): 556–67.

[6] Lewis, "Counterfactual Dependence and Time's Arrow," *Noûs*, XIII, 4 (November 1979): 455–76.

[7] Fred Dretske, "Conclusive Reasons," *Australasian Journal of Philosophy*, XLIX, 1 (May 1971): 1–22; Nozick, *op. cit.*, p. 321.

[8] Jerry Fodor, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge: MIT, 1987).

[9] Risto Hilpinen, "Approximate Truth and Truthlikeness," in M. Przelecki, K. Szaniawski, and R. Wojcicki, eds., *Formal Methods in the Methodology of the Empirical Sciences* (Dordrecht: Reidel, 1976), pp. 19–42.

however, varying from one context to the next; so, he argued, it cannot support the distinctions that philosophers would rest on it.[10] Many specific difficulties have since arisen with philosophy built on comparative overall similarity, despite such misgivings.[11] But let us return to the neighborhood of the observation about weighting properties. There lurks real trouble.

The trouble comes to light when we ask just how to combine similarities and differences in various respects. In fact, no one has had any real idea! There are only metaphors, however promising these might seem. David Lewis draws a comforting analogy to vector addition, with talk of "resultant" similarities.[12] Robert Nozick conjures an image of the judicious balancing of similarities against differences when he speaks of the "weights" of relevant properties.[13] Everyone seems to picture a space, framed by the dimensions of comparison, in which similar things are close together and dissimilar things are far apart.

I shall argue that all these metaphors are false. We cannot add up similarities or weigh them against differences. Nor can we combine them in any other way. Goodman was right to be skeptical. No useful comparisons of overall similarity will result.

My first main point in support of this conclusion will be that there really does have to be a balance of similarities if there are to be useful overall similarities. Greater similarity in one respect will have to make up for less similarity in another respect. Section II asks how to combine similarities and answers in terms of supervenience. Then it considers several ways of combining similarities without weighing them and shows that each fails a reasonable requirement. One of these requirements is that there should not be a *dictator*, that is, a critical respect of similarity that excessively influences overall similarities.

The next main point is that there is no balance of similarities. Section III illustrates the idea of a balance of dimensions with a spatial

[10] Nelson Goodman, "Seven Strictures on Similarity," in Goodman, ed., *Problems and Projects* (Indianapolis: Bobbs-Merrill, 1972), pp. 437–46.

[11] For difficulties with Lewis's treatment of *de re* modality, see Fred Feldman, "Counterparts," this JOURNAL, LXVIII, 13 (July 1, 1971): 406–09; Allen Hazen, "Counterpart-Theoretic Semantics for Modal Logic," this JOURNAL, LXXVI, 6 (June 1979): 319–38; and Michael Fara and Timothy Williamson, "Counterparts and Actuality," *Mind*, CXIV, 453 (January 2005): 1–30. For difficulties with the interpretation of counterfactuals in relation to comparative overall similarities, see Jonathan Bennett, "Counterfactuals and Possible Worlds," *Canadian Journal of Philosophy*, IV, 2 (December 1974): 381–402; Kit Fine, "Critical Notice, *Counterfactuals*. By D. Lewis, Oxford: Blackwell, 1973," *Mind*, LXXXIV, 335 (July 1975): 451–58; and Paul Horwich, *Asymmetries in Time* (Cambridge: MIT, 1987).

[12] Lewis, "Counterpart Theory."

[13] Nozick, *op. cit.*, p. 33.

example. Then it argues that, in general, greater similarity in one respect will not make up for less similarity in another. Similarities are incommensurable when they are merely ordinal and we cannot meaningfully say how much more or less alike things are, but they also are incommensurable when they are cardinal, and we can.

Sections II and III avoid technicality in order to develop an intuitive sense of the trouble with overall similarity. As a result, the argument is less rigorous than you might wish. It considers only a few representative ways of combining similarities. You might wonder whether some other way is better. Also, because the discussion remains informal, there is room for unwanted assumptions to slip in unnoticed. You might wonder whether some such interloper is the real troublemaker.

Section IV gives skepticism about overall similarity a precise sense and a completely rigorous justification. After making matters from the previous sections technically explicit, a reinterpretation of Kenneth Arrow's theorem of social choice shows that a relation of comparative overall similarity must always have a dictator if it supervenes on similarities in several respects.[14]

If all this is so, why then has overall similarity seemed such a promising foundation for philosophy? Perhaps this is because we imagine that our everyday thinking depends on it. For instance, you might have thought that we implicitly compare overall similarities when sorting things into categories. Since we often agree among ourselves about what is what, it is perhaps only natural to suppose that, in categorization, we latch onto genuine overall similarities and differences of things. Certainly, it feels as if we are onto something real.

However, there is another explanation for our agreement. Presumably, there is an innate psychological basis for categorization that does not vary greatly across our species. We are bound to find ourselves agreeing quite a bit, given that we all categorize in much the same way.

Whether the psychological mechanisms of categorization reveal genuine overall similarities is another matter, though, and they need not do so at all. Overall similarities are involved in categorization according to one influential proposal.[15] But things are similar or different, in the relevant sense, only indirectly, through the mediation of mental representations that pick out some features as salient. Such

[14] The relevance of social-choice theory for this topic has gone largely unremarked, but Williamson touches on it in "First-Order Logics for Comparative Similarity," *Notre Dame Journal of Formal Logic*, XXIX, 4 (Fall 1988): 457–81.

[15] See Amos Tversky, "Features of Similarity," *Psychological Review*, LXXXIV, 4 (July 1977): 327–52.

mediated similarities might play a role in our everyday thinking even if things are not in themselves similar or different, independently of how we represent them to ourselves.

Moreover, categorization might not depend on any overall similarities, whether or not they are mediated by representations. When you judge someone to be drunk, for instance, because he has jumped fully clothed into a swimming pool, you need not have done so by establishing an overall likeness to other drunks. Instead, you might have *explained* what happened, drawing on a more-or-less implicit theory of human behavior and the effects on it of too many drinks.[16] However categorization feels from the inside, so to speak, it need not rely on relations of overall similarity among things. Perhaps it is our intuitive sense of similarity and difference that depends on our ability to categorize and not the other way around.

II. THERE MUST BE A BALANCE

I now shall argue that the metaphor of weighing similarities is to be taken quite seriously. Greater similarity in one respect will have to make up for less similarity in another, if there are to be useful overall similarities.

To get started, we will need some understanding of what it is to combine similarities. It should be compatible with adding them up and weighing them against differences as well as with other suggestive metaphors for what is involved: weaving similarities together, or what have you. Fortunately, we can make do with very little understanding. We will proceed with the idea that overall similarities supervene on particular similarities, comparisons of overall similarity being the same whenever all comparisons of particular similarity are the same. In section IV, we will have a completely precise formulation. Meanwhile, an example will illustrate.

Imagine looking over the preserved ship of Theseus on a fine day. In a corner somewhere, you notice the pile of decayed planks that have been removed over the years. You judge, perhaps, that the preserved ship resembles the original ship of Theseus more closely, overall, than the pile does. Soon afterward, back for another look, you find both ship and pile to be just as you left them. Neither has become in any way more like the original ship of Theseus, and neither has become less like it. What supervenience requires is that the overall

[16] Compare Gregory L. Murphy and Douglas L. Medin, "The Role of Theories in Conceptual Coherence," *Psychological Review*, XCII, 3 (July 1985): 289–316, see p. 295. For further discussion, see Ulrike Hahn and Michael Ramscar, eds., *Similarity and Categorization* (New York: Oxford, 2001).

comparison remains unchanged. On both occasions, your judgment ought to be the same.

Do not be misled about the supervenience of overall similarities by the slack between them and their expressions in thought and language. Lewis used the idea that counterfactual conditionals are contextual in order to explain how someone can meaningfully assert either of a contrary pair. Suppose that in an ordinary conversation someone claims:

> If Caesar had been in command [in the Korean war], he would have used the atom bomb.

As soon as the person has spoken, Lewis argues, we rush to help him to have spoken the truth. We evaluate his utterance by using a relation of overall similarity among possible worlds that attaches greater importance to similarity in respect of the knowledge of weapons common to commanders in Korea. With this accommodation, worlds in which Caesar has a modern knowledge of weapons are more similar to our actual world, overall, and the speaker's utterance is true. If, on the other hand, he says:

> If Caesar had been in command, he would have used catapults,

we instead attach greater importance to historical knowledge, and this becomes the true utterance. According to Lewis, we evaluate differently in the two cases because we evaluate with different relations of overall similarity.[17] If he is right about this, the truth of a counterfactual that reveals the overall similarity of worlds need not supervene on their particular similarities, even though, we should suppose, overall similarity itself does supervene.

Consider now some ways of combining similarities without weighing them against differences. Comparing overall similarities is completely straightforward in some rather special cases. We will judge that one person resembles you as closely as another does, overall, if he *dominates*, which is to say that he resembles you as closely in every respect. The overall comparison is easy because there is no need for weighing. It does not depend on how much similarity in the one respect goes for how much dissimilarity in the other—nor on whether there are any such rates of exchange at all.

In cases of dominance, comparative overall similarity is just as transparent and dependable as familiar mathematical notions of similarity, such as congruence among geometrical figures and isomorphism among structures. If overall comparisons can be made

[17] Lewis, *Counterfactuals*, p. 67.

only in such cases, however, then comparative overall similarity will not be of much use in philosophy.

For one thing, you will lack counterparts. Take your spitting image. He walks like you, and he talks like you. He resembles you as closely as can be, except for this: he grew up in a traveling circus. It is hard to imagine a more likely candidate, and yet he will not qualify as your counterpart—not if he lives in any normal possible world and similarity in respect of origins is relevant. Normally, there will be other candidates whose origins are more like yours, and your spitting image will not stand a chance against even the least likely of these. He fails to outdo them in overall likeness to you because, being in the one way less like you than they are, he does not dominate them. It is likely that none of them dominates *him*, either, but we cannot allow that to qualify him as your counterpart. That would make it too easy to qualify. You would wind up with too many counterparts.

In some special cases, then, all comparisons of particular similarity align. One candidate dominates the other and is more similar to you, overall. But it will not do for these to be the only cases in which overall comparisons are available. Something will have to close the comparability gaps.[18]

Your spitting image does not quite dominate the other candidates, but he is not far off. One might suppose that a candidate resembles you as closely, overall, if he *nearly* dominates, that is, if he resembles you as closely in nearly every respect.

This supposition will close comparability gaps because, as with full dominance, what counts is just the proportion of respects of greater similarity. No weighing is called for. Also, it will account for many intuitive similarity judgments. Still, it is not a suitable foundation for philosophy. One complication is the vagueness of "nearly every": just how close to complete agreement among the various dimensions must we come in order for composition to occur? The real problem, though, is that there are incoherent results.

Consider a case patterned on Condorcet's "paradox" of voting. Three candidates compete to be your counterpart. In one respect, Alfie resembles you more closely than Bozo does, and Bozo resembles you more closely than Coco does; in another respect, Bozo is most like you, followed by Coco and then Alfie; and, in some third respect, Coco is most like you, followed by Alfie and then Bozo. Let these dimensions be the only relevant ones:[19]

[18] The standard assumption is that comparative similarities are *connected*: one of any two things resembles you at least as closely as the other one does.

[19] Or, let the candidates resemble you equally in all other respects.

	Increasing Resemblance to You ←
First Respect:	Alfie, Bozo, Coco
Second Respect:	Bozo, Coco, Alfie
Third Respect:	Coco, Alfie, Bozo

Alfie resembles you more closely than Bozo does in every respect but one; so, no matter where the threshold for composition is set, short of full dominance, Alfie nearly dominates Bozo, and Alfie resembles you more closely than Bozo does, overall. Likewise, Bozo resembles you more closely than Coco does, overall. Coherence requires the comparisons to be transitive;[20] in particular, it requires that Alfie resembles you more closely than Coco does, overall; but, by this reckoning, he does not. On the contrary, since Coco resembles you more closely than Alfie does in the second and third respects, the comparison between them comes out the wrong way around. To the extent that our intuitive similarity judgments track near dominance, in a range of possible cases, they are very much the worse for it, because they are incoherent.[21]

For another try at combining similarities, suppose we somehow rank the respects of similarity in order of their importance. Then we can obtain overall comparisons by alphabetic composition, allowing each successive relation of comparative similarity in some respect to refine the result of putting together its predecessors, by breaking ties. This closes comparability gaps, and there is no weighing. For example, more similarity in a lesser respect will never make up for less similarity in a more important one.

You might wonder how to rank the respects of similarity. One idea has been that their relative importance is revealed by our counterfactual judgments. To see how, consider a well-known objection to the theory that the truth of a counterfactual depends on the truth of its consequent in all most-similar antecedent worlds.[22] It easily can be imagined that during the nuclear alert of 1973:

> If Nixon had pressed the button, there would have been a nuclear holocaust.

This is puzzling if it requires that some possible world in which he pressed the button and set off a holocaust is more like our actual world,

[20] The standard assumption is that comparative similarities are *weak orders*—that is, connected and transitive. See Lewis, *Counterfactuals*, p. 48.

[21] Someone might try to save the idea of comparative overall similarities by saying that sometimes composition does not occur and that this is such a case. And he might say that it also does not occur with the profiles in the proof of the theorem in section IV. But saying so would be unwise. There is nothing funny about this case or about the possibilities that those profiles represent. We may expect composition to occur here, if it ever does.

[22] See Fine, *op. cit.*, p. 452; and Bennett, *op. cit.*

overall, than is any world in which he pressed the button but no holocaust followed. You would have thought that, for any given world in which life as we know it was wiped from the face of the Earth, there always will be another, more like our actual world, in which Nixon pressed the button and nothing very much happened. Some miraculous little glitch saved the day. His moment of truth came and went, and he sat there trembling for a good long time. After he pulled himself back together, though, life went on pretty much as it actually did.

Lewis responded to this example by arguing that there are many relations of overall similarity, corresponding to different rankings of the various dimensions, and that the similarities implicit in our counterfactual judgments do not need to be the same ones that our explicit similarity judgments reveal.[23] He then proposed a ranking that he took to be correct insofar as it makes the right conditionals true. In the same vein, Nozick argued that we can discover the ordering of dimensions in our identity judgments.[24] They both had in mind what we might call *revealed* overall similarities, the weights or priorities of dimensions being implicit in which counterfactuals and identities we take to be true and which false.[25] This is how we might hope to rank the respects of similarity. If *priorities* are what we discover, not weights, then the revealed similarities might be alphabetic orders.

However, alphabetic orders are unsuitable no matter how the dimensions are ranked. This is because they are:

> *Dictatorial.* There is a critical respect of similarity such that whenever some things are more similar in this one respect than are some others, their overall similarity is at least as great.

[23] Lewis, "Counterfactual Dependence."

[24] Nozick, *Philosophical Explanations*, pp. 34–35.

[25] Given that the revealed similarities strikingly disagree with our ordinary sense of similarity and difference, you may wonder how we could ever have become attuned to them. As Horwich writes in *Asymmetries in Time* (p. 172):

> [T]hese criteria of similarity might well engender the right result in each case. However, it seems to me problematic that they have no pre-theoretic plausibility and are derived solely from the need to make certain conditionals come out true and others false. For it is quite mysterious why we should have evolved such a baroque notion of counterfactual dependence. Why did we not, for example, base our concept of counterfactual dependence on our ordinary notion of overall similarity?

A further problem with Lewis's ranking is that it in fact does not engender the right results. See Adam Elga, "Statistical Mechanics and the Asymmetry of Counterfactual Dependence," *Philosophy of Science*, LXVIII, 3, Supplement: Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association. Part I: Contributed Papers (September 2001): S313–24; and Barry Loewer, "Counterfactuals and the Second Law," in Huw Price and Richard Corry, eds., *Causation, Physics and the Constitution of Reality* (New York: Oxford, 2007), pp. 293–326.

The critical respect is the one with first priority.[26]

Dictatorship is pernicious. Lewis warned of its excesses: "respects of similarity and difference trade off. If we try too hard for exact similarity…in one respect, we will get excessive differences in some other respect."[27] The problem with dictatorship is precisely that it enforces trying too hard for exact similarity in the critical respect. In the metaphor of balancing, there is no judicious weighing of similarities against differences. Greater similarity in the critical respect simply locks up the balance, preventing it from tipping the other way no matter which differences pile up on the other side. In terms of a similarity space, as things become closer in the critical dimension, they can become only closer overall, no matter how distant they become in other dimensions. Overall similarities are the "resultant" of a multitude of particular similarities and differences only in a tortured sense of the word.

Dictatorship not only offends against the very idea of overall similarity. It also compromises philosophical theories that build on it. Take for instance Lewis's counterpart-theoretic account of *de re* modality. We should expect it not to be committed to the doctrine that things have some of their attributes essentially, independently of how they are specified. Lewis thought that it was.[28] But dictatorship imposes essentialism.

Under dictatorship, there is some critical respect of comparison that trumps the others. Consider any one of your candidate counterparts who is even slightly unlike you in this respect. He will not qualify as your counterpart—not if he shares his world with another candidate who is more like you in the critical respect; under dictatorship, this other candidate must resemble you at least as closely overall.[29] To qualify as your counterpart, a candidate must resemble you in the critical respect at least as closely as his competitors do. Your counterparts are bound to resemble you, in this respect, as closely as can be.

[26] Alphabetic composition actually produces a more severe dictatorship in which things more similar in the critical respect are not merely as similar overall but indeed are more so. I consider the milder sort here because this is the one that returns in section IV.

[27] Lewis, *Counterfactuals*, p. 9.

[28] Lewis writes that "a suitable context might deliver an antiessentialist counterpart relation—one on which anything is a counterpart of anything, and nothing has any essence worth mentioning." See "Postscripts to 'Counterpart Theory and Quantified Modal Logic'," in *Philosophical Papers*, vol. 1 (New York: Oxford, 1983), p. 43. Essentialism is more natural in other accounts of *de re* modality, such as Saul Kripke's in *Naming and Necessity* (Cambridge: Harvard, 1980).

[29] I assume that there is at most one of you in the world in question: you have at most a single counterpart there, who resembles you strictly more closely than all other candidates do.

There is related trouble with counterfactuals. Following Lewis, let us accommodate the claim that Caesar would have used the atomic bomb, by letting similarity in respect of modern knowledge have first importance. Under dictatorship, this respect of similarity is critical. Now consider a possible world in which Caesar's knowledge was completely modern. This world presumably is more similar to our actual world, in the critical respect, than is any world in which his knowledge was not completely modern; so, in the context created by the speaker's claim, it is as similar to our actual world, overall. That is, in this context, no world with an incompletely modernized Caesar is more similar to our actual world, overall, than is the world with the completely modernized Caesar.

Now there is a problem. Intuitively, you can agree with the speaker that Caesar would have used the bomb if he had been in command in Korea, while thinking to yourself that he also, as the need arose, would have used catapults, pila, and other kinds of weapons—even ones that have been forgotten over the millennia:

> If Caesar had been in command, he also would have used long-forgotten weapons.[30]

Under dictatorship, though, you would be mistaken. The truth of this sentence requires that some world in which Caesar was in command and used long-forgotten weapons is more similar to our actual world, overall, than is any world in which he was in command but did not use them. A Caesar who used long-forgotten weapons, though, is an incompletely modernized one. And, as we have seen, when the speaker's utterance is accommodated, no such world is more similar to our actual world, overall, than is the world with the completely modernized Caesar, who did not use long-forgotten weapons because he did not know the first thing about them.

This shows that, just as Lewis warned, trying too hard for exact similarity in one respect can only lead to excessive differences in other respects. Even if one dimension of similarity is most important, the other dimensions still should count as well. There has to be a balance.

We have considered several ways of combining similarities without weighing them against differences, and we have seen that none has a satisfactory result. Composition in the case of dominance is good as far as it goes but leaves too many comparability gaps. Composition in

[30] Be sure to keep this thought to yourself, or you will spoil my example! As soon as you speak up, your accommodating partners in conversation will see to it that you have spoken the truth, by evaluating your utterance in another context, using a different relation of overall similarity.

the case of near dominance fills some of the gaps but has incoherent results. Alphabetic composition imposes dictatorship. There are other ways, but, as we will see in section IV, they are no better. Suitable comparative overall similarities will result, if at all, on the balance of similarities.

III. BUT THERE IS NO BALANCE

I now shall argue that, in general, greater similarity in one respect will not make up for less similarity in another.

It is instructive to contrast similarities with spatial dimensions. Suppose that one person stands closer to you than someone else does. Let him take a single step to the north or south. How far to the east or west should he then move if his relative distance from you is to end up the same as it was to begin with? There is an obvious answer to this question. Any change will do that keeps him on the relevant indifference curve, which, in this case, has a particularly simple form: it is the circle around you that is defined by his starting position. Spatial dimensions are commensurable. In a range of cases, a change in one dimension will make up for a change in another.

We might conceive of corresponding spatial similarities. The overall spatial similarity of two locations, we might say, varies inversely with the great-circle distance between them. It is a function of their similarities in respect of longitude and latitude, which likewise depend on differences in these dimensions. Such spatial similarities are commensurable because the underlying spatial dimensions are commensurable. They inherit their indifference curves from them.

However, the spatial analogy is false. John Maynard Keynes took similarities as an example in making a related point about probabilities:

> [A] book bound in blue morocco is more like a book bound in red morocco than if it were bound in blue calf; and a book bound in red calf is more like the book in red morocco than if it were in blue calf. But there may be no comparison between the degree of similarity which exists between books bound in red morocco and blue morocco, and that which exists between books bound in red morocco and red calf.[31]

The point I take from this is that there is no trading of similarities in respect of the color and the kind of leather of a binding. A book bound in blue morocco bears some overall likeness to a book in red morocco. You can decrease this likeness by changing from morocco to calf, while keeping the color the same. But you cannot regain the original overall likeness to the book in red morocco by subsequently changing the color

[31] John Maynard Keynes, *A Treatise on Probability* (London: Macmillan, 1921), p. 36.

of the calf binding from blue to red. More similarity in respect of color will not make up for less similarity in respect of the kind of leather.

The example involves a dimension that is perhaps merely ordinal. In respect of the kind of leather, the book in blue morocco is more similar to the book in red morocco than is the book in blue calf, but there might be no saying how *much* more similar it is. Perhaps bindings of the same leather are similar in this respect, while those of another sort are different, and that is all there is to it. However, it seems that, in general, there also is no balancing of similarities measured on a cardinal scale.

Take, for instance, similarities in respect of weight and temperature. They might be cardinal, since the underlying quantities are. Let one person resemble you more closely, overall, than someone else does. And let him become a bit less like you in respect of his weight, by gaining a little. Now answer these questions: How much warmer or cooler should he become to restore the original overall comparison? How much more similar in respect of his height? What about his income or his wisdom or hairstyle? That there might be factual answers to these questions is hard to believe.

You might wonder whether the apparent incommensurability is really just a matter of vagueness. True, there is no saying exactly how much more similarity in one respect can be exchanged for less similarity in another, but there might be a rough rate of exchange even so. Indeed, that is just what we should expect, under the assumption that identities, *de re* modal claims, and counterfactuals reveal overall similarities. Normally, what we say is a bit vague. When we utter these sentences, there remain in play several relations of comparative overall similarity, each striking its own balance among similarities and differences. Each relation has a claim to be the right one for the interpretation of what we have said, but none has an exclusive claim. Taken separately, these relations might embody precise rates of exchange among similarities. Taken collectively, they embody none. At best, the context supports rough rates of exchange.

Lewis made a virtue of this vagueness: "[C]omparative similarity is not ill-understood. It is vague—very vague—in a well-understood way. Therefore it is just the sort of primitive that we must use to give a correct analysis of something that is itself undeniably vague."[32] It is the contextual resolution of vagueness that enables him to explain how the contrary counterfactuals about Caesar in Korea can both hold true, each in the context arising from its own utterance.

[32] Lewis, *Counterfactuals*, p. 91.

The problem is that, in general, there do not appear to be rough rates of exchange any more than there are precise ones. Were we aware of any, we should be able to say at least something about them. This we could do by using suitably vague language, as when, before the haggling begins, we can express our rough sense of what things are worth by saying that "only a little" of one will be needed in exchange for a given quantity of another or that, on the contrary, "quite a lot" or "a vast amount" will be needed, as the case may be. When someone has become less like you in respect of his weight, though, we cannot say that he will need to become "only a little" or "quite a lot" or "vastly" more similar in respect of his temperature in order to regain his earlier overall likeness to you. Nor does it seem to be ignorance about the case that keeps us from saying—what could we possibly be missing? As far as we can tell, there are no rates of exchange here.

David Wiggins wrote that values are incommensurable when "there is no general way in which [they] trade off."[33] Similarities seem to be more radically incommensurable than this. There is no general formula for the expression of rates of exchange, such as the circular indifference curves in the case of spatial dimensions. Similarities do not seem to trade off even in highly particular ways, with rates of exchange varying from case to case in complicated ways that defy general description.

Sometimes, perhaps, we should not expect to discover rates of exchange but may make them up to suit ourselves. Consider a speedster built with some salvaged parts. Could it really be "Little Bastard," the very car that James Dean wrecked? If not, how many more original parts would there have to be for the reconstruction to be authentic? How much more causal continuity with the original car would there have to be? Perhaps these are questions for car buffs and their lawyers to settle to their own satisfaction. If this means weighting or prioritizing dimensions and stipulating thresholds for authenticity, perhaps the weights, priorities, and thresholds are theirs to attach and stipulate as they see fit. It is up to them to make up the fact of whether this is "Little Bastard."

But identities, *de re* modal possibilities, and counterfactuals cannot in general depend on made-up similarities.[34] I take it that you are

[33] David Wiggins, "Incommensurability: Four Proposals," in Ruth Chang, ed., *Incommensurability, Incomparability, and Practical Reason* (Cambridge: Harvard, 1997), p. 59.

[34] Even vehicle identities do not, at least for legal purposes. Instead, they depend on the possession of data plates. Basing them on stipulated similarities would not be better. "Little Bastard" then might turn out to be a car, having a temporal boundary coinciding with that of the speedster, but it also might not. Depending on what car buffs and their

unwilling to think that whether something is *you* can be a matter of more or less arbitrary decision or stipulation.[35] It is no easier to accept that some ruling about dimensions, weights, and priorities determines, for example, whether you could have been taller than you are or whether, had it been scratched, the match would have lit. These are not matters that we may settle to suit ourselves.

Admittedly, whether there are rates at which any given dimensions trade off will not be decided in the way that I have approached the matter here, by reflection on how we think and speak. It is an empirical question that, despite all I have said, remains open in the overwhelming majority of cases. In light of this, one might like to think of commensurability as a regulative ideal that guides us toward such rates of exchange as there are to be discovered. Sometimes, indeed, there are surprises: just over a century ago, few could have imagined that spatial and temporal dimensions might be commensurable, but now we know that the temporal order of events is relative to inertial frames and that, as well as spatial indifference curves, there are spatiotemporal ones.[36] Encouraged, one might hold out hope that similarities in different respects will turn out to be commensurable after all.

Time will tell. Meanwhile, the burden will remain on us to discover rates of exchange among similarities for each case separately. For my part, I do not expect that there is much progress to be made in this direction. That is just a hunch, but it also is anybody's hunch that a lot of them await discovery—so many that the idea of a balance of similarities will turn out to be realistic after all. Over the years, a great deal has been built on the notion of comparative overall similarity. The result is an impressive edifice covering large parts of metaphysics and epistemology. Its foundation is about as good as this second hunch.

IV. SIMILARITIES REALLY DO NOT ADD UP

I have argued that there is no good way of combining similarities and differences into useful comparisons of overall similarity. The discussion

---

lawyers decide, and which parts mechanics swap out, "Little Bastard" might come to another sudden end—not with a bang this time but almost imperceptibly when, with the removal of one too many of the original parts, the reconstructed speedster slips below the stipulated threshold for authenticity. In this case, "Little Bastard" will turn out not to have been a car at all but merely an initial temporal part of one, a funny sort of thing like one of Eli Hirsch's "incars" and "outcars" (see "Physical Identity," *The Philosophical Review*, LXXXV, 3 (July 1976): 357–89). This is puzzling: "Little Bastard" was a car if anything ever was.

[35] As Nozick points out in *Philosophical Explanations*, p. 34.

[36] I am told that, in relativity theory, points in space-time at a fixed interval from any given point describe a hyperbola.

was informal, though, and its conclusion remained less than fully secure. I promised a rigorous argument. The first part will make matters from the previous sections technically explicit. Then, a reinterpretation and slight generalization of Arrow's theorem of social choice will show that some respect of similarity always must be a dictator, if comparative overall similarity supervenes on similarities in several respects.

IV.1. Similarities. Comparative similarity is fundamentally a matter of two pairs of things: b resembles b^* as closely as a resembles a^*.[37] The trouble with overall similarity manifests itself in the binary relations that result when b^* and a^* are the same thing—you, for example. For simplicity's sake, we will continue with such relations and with examples having to do with counterparts: $a\mathrm{S}b$ will mean that b resembles you as closely as a does.

We will assume that these relations are *weak orders*:[38]

Connected. For every a and b, either $a\mathrm{S}b$ or $b\mathrm{S}a$;

Transitive. For every a, b, and c, if $a\mathrm{S}b$ and $b\mathrm{S}c$, then $a\mathrm{S}c$.

Connectedness makes the notion of a maximal overall similarity useful (compare the discussion of your spitting image in section II). Given connectedness, coherence requires transitivity as well.

IV.2. Similarity Profiles. These are representations of the similarities and dissimilarities from which comparative overall similarities have been thought to result. One profile concerns your candidate counterparts in one possible world; another concerns the candidates in another world. The domains of profiles may overlap, but they need not do so.[39]

Profiles represent both ordinal and cardinal similarities. Similarities are ordinal when one candidate is more similar to you than another but there is no saying how much more similar. Similarities are cardinal when we can assign proportions to differences—for example, when Alfie and Bozo differ in their resemblance to you twice as much as Coco and Dodo do. There is no need to sort out which similarities are ordinal and which are cardinal, provided that we can accommodate both kinds. No hiding any of the facts from which comparative overall similarity might be thought to result!

Measurement theory has resources for a uniform representation of ordinal and cardinal similarities.[40] Let a *similarity function* be a function

[37] For further discussion of logical aspects, see Williamson, "First-Order Logics."

[38] This is a common assumption. See for example Lewis, *Counterfactuals*, p. 48.

[39] This accommodates the idea that ordinary objects are confined to their own possible worlds.

[40] See for instance Patrick Suppes, "Theory of Measurement," in Edward Craig, ed., *Routledge Encyclopedia of Philosophy* (New York: Routledge, 1998), pp. 243–49.

from some things into real numbers; intuitively, it is a representation of the degree to which these things resemble you, either in some particular respect or overall, as the case may be. Similarity functions are equivalent if they represent the same facts, but what this means depends on whether the facts in question are ordinal or cardinal. One similarity function s is *ordinally equivalent* to another, t, if s is an order-preserving transformation of t, and s is *cardinally equivalent* to t if s is a positive affine transformation of t—that is, there are real numbers $\alpha > 0$ and β such that, for every object o in the domain, $s(o) = \alpha t(o) + \beta$. Here, α allows equivalent functions to use different units, while β makes the origin arbitrary. I assume, then, that any cardinal similarities are to be measured on an interval scale, not on a ratio scale with a fixed origin ($\beta = 0$). This seems right if, unlike mass or heat or other quantities measured on a ratio scale, similarity can neither accumulate nor be entirely absent. This assumption is important, though, and the measurement of similarities will be a good place to start any further investigation into the possibility of aggregating them.

A *similarity measure* is a maximal class of equivalent similarity functions with the same domain. It is ordinal or cardinal, according to the sort of equivalence. Any similarity measure S induces a relation of comparative similarity: $a \check{S} b$ means that, for some (equivalently, all) s within S, $s(a) \leq s(b)$. These induced relations are weak orders. If the domain is assumed to be finite, we may identify ordinal measures with the orders that they induce, but different cardinal measures can induce the same orders. Having distinguished induced orders from similarity measures, I will use 'S' for either and sometimes for both within the same sentence.

Assume there is a (perhaps contextual) finite collection of respects of comparison: 1, ..., n. A *similarity profile* $\vec{S}$ is a list ($\vec{S}_1, \ldots, \vec{S}_n$) of similarity measures, all on the same domain. Each $\vec{S}_i$ is a measure of the similarity to you, with respect to i, of each thing in the domain. The measures of a profile are ordinal or cardinal, according to their respects.

IV.3. Weights and Balance. It is commonly supposed that sometimes one dimension of similarity carries greater weight than another and that it then is possible to combine them. If Alfie resembles you more closely in some respect that carries greater weight, for instance, this supposedly can make up for his resembling you less closely than Bozo does in another respect that carries less weight. Then, on balance, Alfie is more similar to you overall. Whatever it means for dimensions to have weights, presumably things are more favorable for aggregation when they have them. Presumably, comparisons of overall similarity are possible then, if they ever are.

I have argued that often there are no rates of exchange among similarities. Sometimes perhaps there are some, but they are more or less indeterminate. We might suppose that, in general, there are *many* admissible outcomes of aggregation, corresponding to different ways of hypothetically weighting dimensions: the less determinacy there is, the more weightings agree with it, and the more admissible outcomes there are. But reducing indeterminacy to multiplicity in this way does not seem to bring us closer to an understanding of how similarities might add up. I shall now argue that they do not add up even in the most favorable case, in which everything possible has been done to weight them, so that for every profile $\vec{\mathrm{S}}$ there is presumably a unique resultant measure S of comparative overall similarity.

IV.4. Supervenience. I distinguish between two notions. With ordinal supervenience, which of two candidates is more like you, overall, only depends on their comparative similarities in particular respects. With cardinal supervenience, distinctions that are invisible in these ordinal facts may count as well. Ordinal supervenience appears to be the stronger notion, because cardinal facts entail ordinal facts but not the other way around. We will formulate the ordinal assumption and obtain our result. Then, we will see that it still follows when cardinal supervenience is assumed instead.

IV.4.a. To begin, we must capture the idea that, with regard to comparative similarities to you, some candidate counterparts in one possible world are just like some other candidates in another world, in every respect. Let R be a similarity measure; let Δ be some things within the domain of R; and let f be a one-one mapping from Δ into the domain of another similarity measure S. R $\approx_{f,\Delta}$ S means that, for each a and b in Δ, $a\mathrm{R}b$ if and only if $f(a)\mathrm{S}f(b)$. For similarity profiles $\vec{\mathrm{R}}$ and $\vec{\mathrm{S}}$, $\vec{\mathrm{R}} \approx_{f,\Delta} \vec{\mathrm{S}}$ means that, for each i, $\vec{\mathrm{R}}_i \approx_{f,\Delta} \vec{\mathrm{S}}_i$. We assume:

> *Ordinal Supervenience.* For all profiles $\vec{\mathrm{R}}$ and $\vec{\mathrm{S}}$, for all pairs Δ of things from the domain of $\vec{\mathrm{R}}$, and for all one-one mappings f from Δ into the domain of $\vec{\mathrm{S}}$, if $\vec{\mathrm{R}} \approx_{f,\Delta} \vec{\mathrm{S}}$, then R $\approx_{f,\Delta}$ S.

This indicates that which of two candidates is more like you, overall, entirely depends on the ordinal facts of which is more and which is less like you in the relevant respects. Notice two things. First, only the candidates' *similarities* matter: like candidates shall be treated alike, no matter who they are. Second, only *their* similarities matter: they shall be treated alike no matter who else is in the running.[41] There is a further assumption:

[41] Ordinal supervenience is the analogue of Arrow's notion of "Independence of Irrelevant Alternatives," with a slight generalization that allows profiles to have different domains. Arrow named his notion for this second aspect.

Dominance. For every profile $\vec{S}$ and for every a and b in its domain, if $a\vec{S}_i b$ for all i, then aSb.

And here is a definition:

Dictatorship. Among the respects 1, ..., n, there is a critical respect d such that, for every profile $\vec{S}$, if $\alpha\vec{S}_d\gamma$ but not $\gamma\vec{S}_d\alpha$, then $\alpha S\gamma$.

The critical d dictates overall similarities in the sense that, whenever some candidate γ is strictly more similar to you than is another candidate α, in respect of d, γ is at least as similar to you as is α, overall. Now, we have the following:

Theorem. If the similarity profiles and corresponding measures of overall similarity satisfy ordinal supervenience and dominance, then we have a dictatorship.

Proof. See the Technical Annex.

IV.4.b. Allowing overall comparative similarities to depend on cardinal similarities in various respects might be thought to be a way out of trouble, but it is not, if any cardinal similarities are measured on an interval scale.[42]

We will need a notion of cardinal supervenience. Let R be a measure of similarity, let Δ be some things within its domain, and let f be a one-one mapping from Δ into the domain of a measure S. $R \equiv_{f,\Delta} S$ means:

For each $r \in R$, there is some $s \in S$ such that $r|\Delta = s \circ f|\Delta$, and

For each $s \in S$, there is some $r \in R$ such that $s|f(\Delta) = r \circ f^{-1}|f(\Delta)$.

That is, up to the identification of candidates by f, the similarity functions of R, restricted to Δ, are the same as those of S; $\equiv_{f,\Delta}$ generalizes to profiles in the obvious way. Now, instead of ordinal supervenience, we assume:

Cardinal Supervenience. For all profiles $\vec{R}$ and $\vec{S}$ and for all suitable pairs Δ and mappings f, if $\vec{R} \equiv_{f,\Delta} \vec{S}$, then $R \approx_{f,\Delta} S$.

That substitution of cardinal supervenience for the apparently stronger ordinal supervenience is not a way to avoid dictatorship is the point of the following:

Consequence. If the similarity profiles and corresponding measures of overall similarity satisfy cardinal supervenience and dominance, then we have a dictatorship.

[42] Paul Samuelson conjectured that the introduction of cardinal preferences was not a way around Arrow's impossibility theorem. This was verified by, among others, Ehud Kalai and David Schmeidler in "Aggregation Procedure for Cardinal Preferences: A Formulation and Proof of Samuelson's Impossibility Conjecture," *Econometrica*, XLV, 6 (September 1977): 1431–38.

This follows directly from the Theorem and from the fact that, perhaps surprisingly, cardinal supervenience and ordinal supervenience are equivalent. This is because, for all *pairs* Δ (although not in general):

$\vec{R} \equiv_{f,\Delta} \vec{S}$ if and only if $\vec{R} \approx_{f,\Delta} \vec{S}$.

The interesting part is "if." The basic idea of the demonstration is that any two points fall on a straight line and that any two straight lines with the same slope (both up or both down) are positive affine transformations of one another. This means that cardinal similarities, when restricted to pairs, might as well be ordinal similarities. Notice that this is where the assumption comes in that cardinal similarities are measured on interval scales.

MICHAEL MORREAU
University of Maryland at College Park and University of Oslo

TECHNICAL ANNEX: PROOF OF THE THEOREM[43]

An element μ is a *minimum* of relation R if, for each element a of the domain, μRa. Letting R* be the strict relation corresponding to R (xR^*y if xRy but not yRx), μ is a *strict minimum* of R if for each a, $\mu R^* a$. (There are analogous notions of *maxima*.) Take some finite set A with at least three elements and set aside one of them, b. Choose a series of *strict* profiles (all induced relations are strict) on A as follows: $\vec{Q}_0 = (\vec{Q}_{0,1}, \ldots, \vec{Q}_{0,n})$ is any strict profile such that, for each i, b is a minimum of the relation induced by $\vec{Q}_{0,i}$—that is, for every $a \in A$, $b\vec{Q}_{0,i}a$ but not $a\vec{Q}_{0,i}b$. Choose the next profile, $\vec{Q}_1$, such that its induced relations are just like those of $\vec{Q}_0$, except for $\vec{Q}_{1,1}$: b is a strict maximum of this relation.[44] Continuing in this way, we arrive finally at $\vec{Q}_n$; b is a strict maximum of each $\vec{Q}_{n,i}$.

Fact I. Let $\vec{Q}$ be any of the above profiles. Either b is a minimum of the resultant similarity measure Q, or b is a maximum of Q.

[43] This is a slight generalization of one of John Geanakoplos's "Three Brief Proofs of Arrow's Impossibility Theorem," *Economic Theory*, XXVI, 1 (July 2005): 211–15. The only real changes allow profiles to have different domains and to include cardinal as well as ordinal measures. The treatment of cardinal measures comes from Kalai and Schmeidler, *op. cit.*

[44] This construction and another, later one are objectionable, on an intended interpretation. In connection with counterpart theory, profiles represent possible worlds. On pain of begging the question against Lewis's views, we cannot find a series of worlds in which the very same things, represented by the elements of A, are organized differently, since worlds supposedly do not overlap. We can overcome this objection by introducing profile-world isomorphisms. The proof is more easily understood without the added clutter, though; once understood, it is clear enough what is needed.

Demonstration. For contradiction, suppose that b is neither a minimum of Q nor a maximum. Since Q is connected, there are a and c in A, such that cQ^*bQ^*a. Choose another similarity profile, $\vec{P}$, on A such that, for each i, c is ranked strictly above a, while the rankings of each relative to b are the same as in $\vec{Q}$ (simply switch the positions of a and c as needed). Let f be identity. Clearly,

$\vec{P} \approx_{f,\{a,b\}} \vec{Q}$, and

$\vec{P} \approx_{f,\{b,c\}} \vec{Q}$.

Therefore, by ordinal supervenience (where P is the resultant of $\vec{P}$), cP^*bP^*a; so, by transitivity of P, cP^*a. On the other hand, we have chosen $\vec{P}$ so that $a\vec{P}_ic$, for every i; by dominance, this delivers aPc. This is a contradiction.□

Fact II. There is some critical respect of comparison d such that b is a minimum of Q_{d-1}, and b is a maximum of Q_d.

Demonstration. By dominance, b is a minimum of Q_0 and a maximum of Q_n. Suppose there is some $d > 0$ such that b is a minimum of Q_{d-1} but not a minimum of Q_d. By Fact I, b is a maximum of Q_d. Otherwise, let $d=n$.□

We now will see that the critical respect of Fact II is a dictator: for any profile $\vec{S}$ and for any objects α and γ in its domain,

if $\alpha\vec{S}_d{}^*\gamma$, then $\alpha S\gamma$.

To this end, suppose $\alpha\vec{S}_d{}^*\gamma$. Let B be the domain of $\vec{S}$, and choose anything β that is not in B. Choose another profile $\vec{R}$ whose domain is $B \cup \{\beta\}$, such that the induced relations $\vec{R}_i$ satisfy, for all $\delta, \varepsilon \in B$:

For every i, $\delta\vec{R}_i\varepsilon$ if and only if $\delta\vec{S}_i\varepsilon$;

for every $i < d$, $\delta\vec{R}_i{}^*\beta$;

$\alpha\vec{R}_d{}^*\beta$ and $\beta\vec{R}_d{}^*\gamma$; and

for every $i > d$, $\beta\vec{R}_i{}^*\delta$.

$\vec{R}$ is just like $\vec{S}$, except that β ranks strictly above everything else in the orders before the critical dth order, between α and γ in this order, and below everything else in the remaining orders. Let f be identity, and note first that:

(1) $\vec{R} \approx_{f,\{\alpha,\gamma\}} \vec{S}$.

Let a be any element of A other than b, and let g be a mapping such that $g(a) = \alpha$ and $g(b) = \beta$; note also that:

(2) $\vec{Q}_d \approx_{g,\{a,b\}} \vec{R}$.

Let c be any element of A other than b, and let h be a mapping such that $h(b) = \beta$ and $h(c) = \gamma$; note also that:

(3) $\vec{\mathrm{Q}}_{d-1} \approx_{h,\{b,c\}} \vec{\mathrm{R}}$.

By Fact II, $a\mathrm{Q}_d b$ and $b\mathrm{Q}_{d-1}c$. By (2) and (3) above and by ordinal supervenience, $\alpha\mathrm{R}\beta$ and $\beta\mathrm{R}\gamma$; so, by transitivity, $\alpha\mathrm{R}\gamma$. Finally, by (1) and ordinal supervenience, $\alpha\mathrm{S}\gamma$. This completes the proof.

A NOTE ON PARITY AND MODALITY*

Towards the end of his *Naming and Necessity*, Kripke raises the question of whether every physical necessity is a necessity *tout court*. Responses to this question are, schematically, of three kinds. There are those who hold that there is no such thing as "merely" physical necessity, as opposed to absolute metaphysical necessity, thereby expanding the kind of necessity that Kripke ascribes to "Water is H_2O" to all the laws of nature. Others claim that they can imagine the constants of nature being different, or that it is possible to conceive laws which are inconsistent with the laws of nature.[1] The third response is an important locus of disagreement between Kripke and Putnam,[2] who holds views similar to Kripke's on the nonepistemic nature of physical necessity, but requires that on the issue of metaphysical necessity we must come closer to the empiricist. In this paper I wish to argue for a possibility (or necessity) that is not physical but still not merely logical or mathematical. Unlike discussions which try to answer our question by offering a comprehensive account of the laws of nature and touch on general issues such as properties and identity, I wish to show the richness and possible fruitfulness of studying one specific example: the breaking of left and right symmetry in weak interactions. Some of the problems we are likely to encounter in suggesting a possibility that is not consistent with the laws of nature do not arise when considering such symmetry breaking at the fundamental level.

I. THESIS

The weak force (or: 'weak interactions') is one of the four fundamental forces in nature. Among its manifestations is beta decay. The topic

*Dedicated with thanks to Hilary Putnam on the occasion of his receiving the Prometheus award.

[1] This division to camps could result in oversimplification, as each one may use a different notion of metaphysical necessity. In this note there is no place to survey the vast literature on the subject that has emerged since *Naming and Necessity*. See for example E. Jonathan Lowe, "Kinds, Essence, and Natural Necessity," in Andrea Bottani, Massimiliano Carrara, and Pierdaniele Giaretta, eds., *Individuals, Essence and Identity: Themes of Analytic Metaphysics* (Boston: Kluwer, 2002), pp. 189–206; John Bigelow, Brian Ellis, and Caroline Lierse, "The World as One of a Kind: Natural Necessity and Laws of Nature," *British Journal for the Philosophy of Science*, XLIII, 3 (September 1992): 371–88; Chris Swoyer, "The Nature of Natural Laws," *Australasian Journal of Philosophy*, LX, 3 (1982): 203–23, at p. 214; Crawford L. Elder, "Laws, Natures, and Contingent Necessities," *Philosophy and Phenomenological Research*, LIV, 3 (September 1994): 649–67.

[2] Hilary Putnam, "Is Water Necessarily H_2O?" in *Realism with a Human Face* (Cambridge: Harvard, 1990), pp. 54–79.

0022-362X/10/0709/491–498

can be approached at greater or lesser depth, but to understand the main point of this note a popular exposition will do. Beta decay is a form of radiation that was discovered in 1896 by Henri Becquerel and named in 1899 by Ernest Rutherford. Yet its nature remained enigmatic until it was clarified by Enrico Fermi in 1934. According to Fermi, the electron (or positron, if it is a beta plus process) which emerges from the nucleus is the result of a decay of a neutron into a proton, an electron, and an anti-neutrino:

$$n^0 \rightarrow p^+ + e^- + \bar{\nu}_e$$

There are good arguments which show that this decay results from a force that cannot be identified with known forces. It was named the "weak force" because it was found to be much weaker than the strong and the electromagnetic forces.

One of the most astonishing discoveries about weak interactions occurred more than twenty years after Fermi's clarification, when Chien-Shiung Wu confirmed a prediction made by the two young physicists Chen-Ning Yang and Tsung-Dao Lee, and found that the weak force is not symmetric with respect to reflections.[3] Working with decays of neutrons from heavy atoms (cobalt 60), she was able to show that most of the electrons that are emitted in this process prefer a direction that is opposite to the magnetic moment of the nucleus. To explain this in the simplest form we may use the image of umbrellas. Umbrellas at rest can be thought of as congruent to each other.[4] But once they start to spin you will have to divide them into two equivalence classes. What Wu discovered is analogous to Nature preferring one kind of rotating umbrellas—left-handed ones—over the other. When a molecule rotates, electrons can be emitted in either of two directions that are perpendicular to the plane of the rotation. Yet *most* of the electrons move in one specific direction.

Weak interactions are unique in that they break the symmetry of mirror reflection. Thus, while with the other three forces reflection

[3] C.-S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental Test of Parity Conservation in Beta Decay," *Physical Review*, CV, 4 (1957): 1413–15.

[4] Oliver Pooley's note is appropriate here: "It is worth stressing that the relevant notion of possibility here is not that of physical possibility. It is physically impossible to superpose a left hand and its perfect *left-handed* duplicate if they are both solid material objects. Rather we must abstract from such physical limitations and consider whether it is *mathematically* possible for the distances between the two objects to be changed continuously in such a way that the two objects eventually coincide." See Pooley, "Handedness, Parity Violation, and the Reality of Space," in Katherine Brading and Elena Castellani, eds., *Symmetries in Physics: Philosophical Reflections* (New York: Cambridge, 2003), pp. 250–80.

is possible for every interaction, it is a law of nature that reflection of the weak force is impossible in the beta decay. This is certainly related to Kant's argument for his transcendental philosophy from the incongruity of the right and left gloves. Indeed, the violation of the reflection symmetry has provoked a lively debate on the validity of Kant's argument. My aim, however, is only to show the relevance of this subject for conceiving a metaphysical possibility that is not a physical one. Thus: although it is physically impossible for a weak interaction to prefer the right hand over the left, a world with such a preference is possible. Further, this suggestion is rich enough to withstand—at least in the first round—four of the problems commonly raised against metaphysical possibilities.

II. FIRST QUESTION

How Do You Know That This Is a Real Possibility? Let me start by saying that the possibility of a world in which the weak force prefers the right hand over the left one seems to me rather intuitive. I am thinking of a mirror world in which every event (x, y, z, t) is transformed to $(-x, -y, -z, t)$. Every inconsistency in the mirror world can be translated to a problem in our world, but since our world is possible the mirror world must also be possible.

A person who says that it is possible for water to boil at 200 C (in normal conditions), or that it is possible for Ohm's law to take a different form, is usually answered, and rightly, that he may not have given attention to the fact that other parts of physical theory—other fundamental laws—are in contradiction with these possibilities. In such examples, one focuses on part of the world and theory and fails to notice a special problem elsewhere. But our mirror world is a *complete* world, and our intuition is that we are not overlooking a problem somewhere, for we are speaking of a whole world. In this respect it is different from imagining water boiling at 200 C, and since we are discussing a fundamental principle there is no danger that we might stumble here as we do in the case of Ohm's law.[5]

One of Kant's insights may support the point I am making here from another angle. The mathematical world does not tell us anything about right and left. Numbers, functions, equations, and operators can be asymmetric, but nothing in this asymmetry can be identified with the right and left hand in nature. The basic laws of nature are written in terms of differential equations and Lagrangians: they give us the form of the phenomena, and thus, although they

[5] More needs to be said if we do not take this as a fundamental law; see discussion of the third question below.

are applied to the real world, to space and time, the mathematical formulations of laws cannot tell you whether this hand or that hand is right. The latter, to borrow terms from Kant, is *shown*, not *said*. Even if laws dictate an asymmetry, they cannot tell us how it should be placed in the world.[6] This gives rise to the idea that there is some degree of freedom in the passage from the laws of nature to the world. The idea that a mirror world is possible relies on the gap between the pure world of mathematics and our ordinary world with its objects.[7]

True, one may invoke the law that is true in our world to claim that my mirror world only looks like a possible world, but in my view *this* objection begs the question. The intuition to which I am appealing may be refutable, but not by repeating the law that weak interactions prefer the left hand in our world. I hope that the following discussion will add to what I have explained above.

III. SECOND QUESTION

Can We Separate the Possibility We Are Entertaining Here from a Logical One? Or, More Strongly, Are We Discussing a Possibility That Is about Weak Interactions at All? I believe that we should answer these questions in the affirmative. The weak interaction was studied through experiments on beta decay, and we did not think of the preference of the left hand as part of the definition or essential characteristic of this kind of interaction; in fact, we expected weak interactions to behave like all the other forces and be symmetric with respect to reflection. We were astonished to discover that the situation is different.

Now, one may resist this suggestion by saying that while the asymmetry was not part of the essence of the concept of weak interaction in 1899, today we have discovered that this is the case. But physicists do not speak this way. Some even think that we may be wrong and that weak interactions somehow really are symmetric with respect to reflection. In considering these possibilities they continue to use the term "weak interaction." Moreover, we believe that weak interactions are carried by bosons, W+, W−, and Z. When we read the reason for the violation of the reflection symmetry from the Lagrangean we describe it as the interaction of the bosons W+, W−, and Z with leptons and quarks.[8] The unactualized possibility we are contemplating

[6] Note that in suggesting this possibility we are not changing the physical theory. Below I shall mention the possibility that the Z, W+, and W− bosons interact differently with leptons, which yields a different way to argue for the possibility of a mirror world. The two ways have different assumptions and should be kept separate.

[7] This must be connected to the question of the reality of space-time. I am, however, not sure that this is tantamount to holding a nonrelationist theory of space.

[8] A semi-popular explanation is given in Roger Penrose, *The Road to Reality: A Complete Guide to the Laws of the Universe* (New York: A. A. Knopf, 2004), pp. 628–45.

here would involve a different kind of interaction between bosons and leptons. Since these bosons are the carriers of the force we are allowed to claim that in our hypothetical case we are talking about weak interactions.

Indeed, logical possibility is the lack of contradiction. When we say "Water is not H_2O" is logically possible, we can either show syntactically that we can never derive a contradiction from this assumption, or we can create an interpretation of the terms in this sentence such that it will make the sentence true. In making this interpretation we are free to choose any object we wish, since the consistency of our sentence has nothing to do with either water or atoms.[9] Yet, when we are considering asymmetry in weak interactions, we are specifically discussing electrons, neutrons, and so on, and not any other possible object. We therefore should separate the possibility we are entertaining here from a logical one.

IV. THIRD QUESTION

Can an Explanation Eliminate the Asymmetry? Here the suggestion is that further developments and discoveries may force us to get rid of this asymmetry. I assume that the violation of symmetry by weak interactions is a genuine fact. Unlike the case of Orsted's needle, this is not a breach of symmetry that we may attribute to having neglected to observe the behavior of electrons and other particles. Now, to fully explain an asymmetry, and not only to explain it away, we must assume another asymmetry. Pasteur's explanation of the difference in our reaction to symmetric molecules by assuming chirality in the molecules that constitute our bodies is one example of this principle, and if we succeed in explaining chirality in organic molecules by reducing it to the asymmetry in weak interactions we would have an impressive demonstration of the same principle. This natural principle should be the reply to anyone who questions our view by suggesting that when we accept the possibility of a weak force that prefers right over left it is because of our ignorance of the nature of the weak interactions. Every attempt to explain the asymmetry will only move it one step backward and will not manage to eliminate it.

Allow me to expand on this issue from two complementary angles. If you have a set of laws where *a* and *b* are symmetric, then every

[9] Truth in all possible worlds is different from truth in all interpretations. If 2 to the power of $\aleph_0$ equals $\aleph_1$ then I assume it is true in all possible worlds, but there are many interpretations of set theory in which it is false. The same is true for the sentence "Cicero is Tully." That is why, unlike other writers, I prefer to separate metaphysical possibility from a merely logical one.

deduction from the set of laws has a dual derivation, and the set of all deductions from these sets is also symmetric. Here is the simplest example:

$$R(a,b) \text{ iff } R(b,a)$$

From this we can deduce:

$$\exists x((Rx,b) \text{ iff } R(b,a))$$

Which is not symmetric. But there is a different conclusion that "saves" the symmetry, its dual one:

$$\exists x(R(x,a) \text{ iff } R(a,b)$$

Thus, if by explanation of the asymmetry in weak interactions we mean the derivation of the preference of the right hand, then the dual derivation from symmetric principles would show an opposite preference. This shows that if explanation is a derivation then we cannot explain asymmetric laws from symmetric ones.

Second, sometimes we have processes that spontaneously break symmetry. A ferromagnet, as is discussed in many textbooks on symmetry breaking, can be the result of a symmetry breaking in which we pick up a specific direction in space. Such processes are described as a choice of an asymmetric solution to a symmetric differential equation while holding that the set of all solutions retains the original symmetry. This shows that the symmetric differential equation cannot serve as an explanation for the asymmetric situation. Such an explanation will not distinguish between the actual state of affairs and a possible incompatible state.[10] An indication of this can be found in the fact that such processes are not deterministic, and probability considerations enter here in a natural way. These last two considerations support my assumption that symmetry can never explain an asymmetry.

V. FOURTH QUESTION

The possibility of a mirror world can be attacked, however, from a different angle. One might claim, "Surely, the mirror world is possible simply because it is identical with the actual world?!" Behind

[10] The reader who is not familiar with differential equations may think of this simple analogue. The equation $x^2 + y^2 = 5$ is symmetric with respect to x and y. The ordered pair (1,2) is an asymmetric solution, but so is (2,1). The symmetry in the original equation is preserved in the set of the solutions. Every asymmetric solution has a "dual" one which is incompatible with it. Here is a nice suggestion that I find hard to resist: Wherever we move from a symmetric differential equation to an asymmetric solution we introduce contingency into nature.

this question lies a kind of Leibnizian view of space. Much in the same way that it is nonsensical, in a relationalist theory of space, to think of a world where everything in space is transformed one meter to the right, it is meaningless to think of a different mirror world.

If we adopt a nonrelationalist view of space—a stance that is welcomed by the very discovery of the asymmetry in weak interactions—we may, perhaps easily, escape the problem posed by this question. I cannot enter into the subtleties that considering the weak force would have on Kant's view of space. However, I would like to take advantage of this question to add one concluding yet crucial note.

There are at least two ways to speak of left-handedness in other possible worlds. In the first, when we have a possible world we can determine which hand is left by observing which hand is preferred by weak interactions. If this is the way we are talking of left hands in possible worlds, then there is no room for a world where the weak force will prefer the right hand. It is worth noting that this unhappy conclusion is independent of what we think is the nature of space and time. A different way of speaking about handedness in possible worlds, however, is to return to Kripke, who understands reference to a left hand in other possible worlds as a way of speaking about *this* left hand in hypothetical situations. In other words, speaking about left-handedness in other possible words in fact is speaking about left-handed objects that exist in *our* world. In this conception the mirror world is completely different from ours. The difference between the two ways of speaking of left and right in possible worlds is analogous to speaking of south and north in other possible worlds. One may claim that north is defined as the direction of the needle of a compass. On this view, the possibility of the needle pointing to the south is unthinkable. On this view, again, it is impossible to think of a possibility where the needle of the compass will point to the south. But this is not the only way: another way, which seems to me more intuitive, is to point to the north and claim that it is possible that the needle of the compass will point to this direction, and then point to what for us is south. Kripke's question which opened this note therefore is best answered by using his way of referring to objects in other possible worlds.[11]

Let me summarize. There is prima facie evidence that a mirror world is a genuine possibility, neither logical nor physical, that does not owe its reality to our ignorance. To make this statement it seems

[11] Can we answer the fourth question without committing ourselves to a Kripkean way of speaking of possible worlds? I leave this question to a further study.

that we have to assume that explanations preserve symmetries, and adopt a Kripkean view on referring to objects in possible worlds. In this note I could not study general issues such as identity, properties, and so on, but I do believe that a strong connection between identity and laws of nature which does not allow for a possibility that is inconsistent with physical laws requires a reservation.

MEIR BUZAGLO

The Hebrew University

Graduate Faculty Philosophy Journal

The ***Graduate Faculty Philosophy Journal***, published biannualy in association with the Department of Philosophy, New School for Social Research, publishes special issues dedicated to investigating the historical development of concepts and problems in the history of philosophy and their ongoing relevance to issues in contemporary philosophy.

Special Issue

Philosophy, Language, Literature

Issue 31:1 Spring 2010

Contact: Editor, ***Graduate Faculty Philosophy Journal***, Department of Philosophy, New School for Social Research, 79 Fifth Avenue, Suite 1028, New York, NY 10003, email: gfpj@newschool.edu; Electronic and print subscriptions to the GFPJ are handled by Philosophy Documentation Center. Print subscriptions: $25/year for individuals; $15/year for students; $60/year for institutions. Electronic subscriptions: $60/year for individuals; $165/year for institutions. Joint electronic and print subscription: $198/year for institutions. For international orders, please add $8 for shipping. Further subscription information: http://secure.pdcnet.org. The GFPJ is online at **www.newschool.edu/nssr/GFPJ**.

REVUE INTERNATIONALE DE PHILOSOPHIE

Direction: Michel Meyer

c / o Université Libre de Bruxelles, C.P. 188
50, avenue Franklin-Roosevelt / B-1050 Bruxelles (Belgique)
Tél. 32 2 650 25 61 / Fax 32 2 650 45 61
e-mail : tsalis@ulb.ac.be

3-2010 / volume 64 / n° 253

BUDDHIST PHILOSOPHY

COMPTES-RENDUS/BOOK REVIEWS

Prochains fascicules / Forthcoming

Motion Picture
Le théâtre du romantisme à nos jours
Stanley Cavell
Zizek