

AN INCONSISTENCY IN DIRECT REFERENCE THEORY*

My goal is to expose a series of internal difficulties in direct reference theory, ultimately revealing an inconsistency which concerns neither meaning nor the propositional attitudes but rather our standard logic of identity itself. I will not focus on specific formulations, such as those put forward at one time or another by David Kaplan, John Perry, Jon Barwise, Nathan Salmon, Scott Soames, Mark Richard, Mark Crimmins, and many others.¹ By the end it should nonetheless be evident that direct reference theory's basic tenets make this inconsistency inevitable.

I. BASIC TENETS

The story begins with two background tenets accepted by direct reference theorists—and, indeed, nearly all contemporary philosophers of language. First, the expressions 'assert', 'mean', 'know', 'know a priori', 'is necessary', and so forth often function as predicative expressions, and 'that'-clauses, when used in combination with such uses of these expressions (for example, ' $A = B$ ' means that $A = B$; Frege knew that $A = B$), function as singular terms denoting propositions—specifically, the proposition expressed by the sentence embedded within the 'that'-clause. Second, propositions are fine-grained. de Morgan's equivalences (for example, between $A \& B$ and $\text{not}(\text{not } A \text{ or not } B)$) provide a nice illustration of necessarily equivalent propositions that are nevertheless distinct. Direct reference theorists typically deal with fine-grainedness by treating propositions as structured entities, which they identify (or represent) with ordered sets, sequences, or labeled trees.

Contemporary direct reference theory may be understood as a response to the Saul Kripke-Hilary Putnam critique of Fregean semantics—best illustrated by their attack on the Fregean solution to Frege's

* I wish to thank Devon Belcher, Chad Carmichael, Michael Della Rocca, Daniel Korman, Karson Kovakovich, Iain Martel, Marc Moffett, Michael Peirce, and David Robb.

¹ I do not include Ruth Marcus on this list because her views on propositions and the attitudes differ sharply from those of the others and would require a different style of response—see, for example, her "Rationality and Believing the Impossible," this JOURNAL, LXXV, 6 (June 1983): 321–37. I myself defended a form of direct reference theory (in "Semantics" and "Pragmatics," *Quality and Concept* (New York: Oxford 1982) pp. 157–66 and 166–76) but have since abandoned it in favor of a *semantical* account ("Propositions," *Mind*, CVII, 425 (January 1998): 1–32).

Puzzle itself. As formulated by Alonzo Church, Frege's Puzzle is this: How can ' $A = B$ ', if true, differ in meaning from ' $A = A$ '?² Frege's answer is that the difference results from the fact that, even though ' A ' and ' B ' have the same *reference*, they have different *senses*. But arguments by Ruth Marcus, Keith Donnellan, Kripke, Kaplan, Putnam and others have led direct reference theorists to conclude: (i) there are no Fregean senses mediating between a name and its reference; (ii) names do not contribute anything to the meaning of a sentence over and above their reference, and (iii) the reference of a name is in this sense *direct*. Against this background, direct reference theorists attempt to "solve" Frege's Puzzle by just denying its central premise that ' $A = A$ ' and ' $A = B$ ' really can (if true) differ in meaning. On the contrary, if ' A ' and ' B ' refer to the same thing, ' $A = A$ ' and ' $A = B$ ' must also mean the same thing. Call this *tenet I*.

Besides the original semantical version of Frege's Puzzle, there is also a logical version: how, if $A = B$, can the proposition that $A = A$ and the proposition that $A = B$ be different? Direct reference theory may also be understood as a response to the failure of the Fregean solution to this puzzle. More generally, it may be understood as a response to the failure of Fregean logical theory—specifically, its theory of propositions (*gedanke*, that is, thoughts). According to this theory, the constituents of a proposition are not the *objects* the proposition is about but rather modes of presentation (*Artes des Gegebenseins*) of those objects, where these modes of presentation are none other than Fregean senses. So, for example, the third constituent of the proposition that $A = A$ is not the object A but rather the sense of the name ' A ' (for example, the concept of being A), and the corresponding constituent of the proposition that $A = B$ is not the object B (which just is the object A) but rather the sense of the name ' B ' (for example, the concept of being B). The two propositions differ because these two senses differ (the concept of being $A \neq$ the concept of being B). So goes the Fregean solution to the logical version of Frege's Puzzle. But the arguments mentioned earlier (if correct) also show that there can be no plausible Fregean modes of presentation to serve as the relevant constituents of such propositions. This leads direct reference theorists to conclude that the constituents of a proposition are the very objects the proposition is about; in other

² Alonzo Church, "The Need for Abstract Entities," *American Academy of Arts and Sciences Proceedings*, LXXX (1951): 100–13. In what follows ' A ' and ' B ' are to be understood as proper names or common nouns. For convenience, I will often use single quotation marks where corner quotation marks are strictly called for.

words, all propositions are pure *Russellian* propositions. Accordingly, these theorists are once again led to "solve" the puzzle by just denying its central premise, that the propositions in question are in fact different. Instead, they hold the following tenet: if *A* and *B* are identical, the proposition that $A = A$ and the proposition that $A = B$ are identical.³ Call this *tenet II*.

In fact, given the above background tenets, the semantical and logical tenets turn out to be equivalent. Given the disquotational truisms that ' $A = A$ ' means that $A = A$, and ' $A = B$ ' means that $A = B$, it follows that ' $A = A$ ' and ' $A = B$ ' mean the same thing iff the 'that'-clauses occurring in these truisms denote the same propositions. In other words, ' $A = A$ ' and ' $A = B$ ' mean the same thing iff the proposition that $A = A$ and the proposition that $A = B$ are identical. Thus, the consequents of tenet I and tenet II are equivalent. At the same time, we know (by disquotation) that '*A*' and '*B*' refer to the same thing iff $A = B$. Thus, the antecedents of tenets I and II are also equivalent. Hence, the two tenets are themselves equivalent. In what follows, it will be convenient to focus primarily on the logical tenet.

II. PRIMA FACIE CONTRADICTIONS AND THE STANDARD THREE-PLACE RESPONSE

The above tenets, however, leave direct reference theories with *prima facie* contradictions, most famously exhibited by the propositional attitudes. Nearly everyone agrees that the following at least seems intuitively obvious:

- (1) It is not possible to know a priori that Hesperus = Phosphorus.

But tenet II seems to underwrite valid arguments like the following:

- (2) The proposition that Hesperus = Hesperus is identical to the proposition that Hesperus = Phosphorus.
- (3) It is possible to know a priori that Hesperus = Hesperus.⁴

³ A few direct reference theorists deny that these propositions are identical, not because they involve different Fregean modes of presentation, but because they differ in logical form. Since there are other counterintuitive proposition-identities (for example, that Hesperus = Venus and that Phosphorus = Venus) which direct reference theorists cannot avoid by this maneuver, it will be harmless to stick with the simple formulation in the text. I will also assume, as uniformity dictates, that tenet II generalizes from identity propositions to more complex propositions; *mutatis mutandis* for tenet I.

⁴ As is customary, I am suppressing the complication having to do with existence assumptions. Also, I use the Hesperus/Phosphorus example only because of its simplicity; some people might prefer to use a different example.

- (4) Therefore, it is possible to know a priori that Hesperus = Phosphorus.

(2) is an immediate consequence of tenet II plus the well-known astronomical identity. (3) is a trivial fact of our background epistemology. And (4) follows directly from (2) and (3). But (4) contradicts (1). Direct reference theorists rightly take it as a condition of adequacy that their theory explain, or explain away, this and kindred prima facie contradictions. After all, it would not be satisfactory simply to dismiss our intuition about (1), for the aforementioned arguments against Fregean semantics and logical theory are themselves ultimately founded almost entirely on intuitions (modal intuitions concerning Aristotle and the teacher of Alexander, Gödel and Schmidt, twin earth, and so forth). If a relevant difference cannot be found between these intuitions and our intuition about (1)—and kindred intuitions—the case for direct reference theory would be undermined.⁵

Direct reference theorists try to meet this challenge by offering an explanation that in one way or another relies on pragmatic factors (of the sort encountered in Gricean conversational pragmatics). On a representative explanation of this general type, the contradiction is explained away by the pragmatic idea that—although (1) is literally false—when we assertively utter (1) in conversation, the true thing we would be saying (that is, conversationally asserting) would be something like:

- (1') It is not possible to know a priori that Hesperus = Phosphorus
under some mode of presentation similar to 'Hesperus = Phosphorus'.

Here 'know a priori' does not function as a two-place predicative expression 'Know-a-priori² (x, p)' but rather as a three-place predicative expression 'Know-a-priori³ (x, p, m)'.⁶ This paraphrase avoids the

⁵ Some people seem not to accept that this is the dialectical situation but rather think that the evidence for theses I and II is conclusive (and so believe that the propositional-attitude puzzles pose no special threat, being simply unsolved puzzles with which everybody must contend). But this cannot be right. For there exist promising alternative theories that can accept at face value virtually all of the intuitions used to support theses I and II *as well as* the preponderance of our propositional-attitude intuitions. Direct reference theorists thus have no choice but to explain, or explain away, the latter intuitions and then to try to win the larger competition on grounds of overall simplicity and plausibility.

⁶ Although 'a priori' is really functioning as an adverbial phrase in our example, the discussion will be simplified, without bias, by treating 'knows a priori' as a single predicative expression.

contradiction.⁷ First, (1') clearly does not contradict (4): modes of presentation are not even mentioned in (4). Second, it is true that (1') does contradict what we would be asserting conversationally if we were to utter (4) in conversation:

(4') It is possible to know a priori that Hesperus = Phosphorus *under some mode of presentation similar to 'Hesperus = Phosphorus'*.

But this contradiction is irrelevant because (4') simply does not follow from (2) and (3) taken literally, nor does (4') follow from what would be asserted conversationally by uttering (2) and (3). So the contradiction is explained away.⁸

Most direct reference theorists adopt some version of the above three-place explanation. But before we come to our ultimate goal—uncovering the inconsistency between direct reference theory and our standard logic of identity—we must first expose the inadequacy of two familiar two-place accounts of conversational uses of 'knows a priori' and other attitude predicates: (a) two-place descriptive para-

⁷ On an alternative proposal, direct reference theorists might hold that what we understand and intuit when (1) is asserted is really false (and really does lead to contradiction) but then try to account for the source of our error by treating intuiting as a ternary relation and by characterizing the state that we are in when (1) strikes us as true as follows: we intuit *under a mode of presentation similar to 'It is not possible to know a priori that Hesperus = Phosphorus'* that it is not possible to know a priori that Hesperus = Phosphorus. There are two problems with this account. First, unlike the account in the text, this account fails to meet the dialectical challenge. What is called for is an account that does not have the effect of casting doubt on the general reliability of intuition—including, in particular, the sort of intuitions relied upon evidentially by Marcus, Kripke, et al. If false propositions can, as this account suggests, so easily appear true under certain modes of presentation, why trust our intuitions in these evidentially crucial cases? (Kripke, for example, was well aware of this dialectical requirement.) Stalemate and/or skepticism threatens. The second problem is that the present account is in a way incoherent. Since intuiting is being treated as a ternary relation, it would be absurdly ad hoc not to treat other propositional attitudes analogously. The simplest, most straightforward ternary treatment of knowing-a-priori is just the ternary treatment in the text, on which (1') is identified as the propositional content of the intuition we have when an assertion of sentence (1) strikes us as intuitively obvious. But (1') is true, not false, contrary to the account under consideration. So this ternary account of (1) should be incorporated into the present account: given that (1') is true, this emended account would then satisfy the aforementioned dialectical requirement.

⁸ As Stephen Schiffer succinctly puts it "Knowing is either [treated as] a *triadic* relation that holds among a knower, a Russellian proposition the knower knows, and a mode of presentation under which the knower knows the proposition, or else it is [treated as] a *dyadic* relation that holds between a knower and the Russellian proposition she knows, where this dyadic relation is explicable in terms of a triadic relation, $KNO(x, p, m)$, which behaves just as knowing would behave if it were a triadic relation." See "Amazing Knowledge," this JOURNAL, XCIX, 4 (April 2002): 200–02, on p. 200.

phrases and (b) reliance on ordered pairs $\langle p, m \rangle$. The failure of these two-place treatments will force direct reference theorists to adopt the three-place treatment (or a certain nonstandard two-place treatment in which 'under mode m ' is a sentential or adverbial operator), and doing so will, in turn, force them into the advertised inconsistency.

III. TWO-PLACE DESCRIPTIVE PARAPHRASES

The first of these two-place accounts is reminiscent of Kripke's well-known descriptivist paraphrases of sentences such as 'It could have turned out either way whether Hesperus was Phosphorus'.⁹ On this sort of account, when one assertively utters sentence (1) in conversation, the true thing one would be saying is:

It is not possible to know a priori that the first heavenly body visible in the evening = the last heavenly body visible in the morning.¹⁰

The same idea may be used to resolve related prima facie contradictions. Suppose you know virtually nothing about the planets beyond the trivial truth that Hesperus = Hesperus. Suppose, however, that after reading a pamphlet on the solar system you report: I just learned from the pamphlet that Hesperus = Phosphorus. (Here and below 'learn' is used in its factive sense to mean coming to know.) But, according to direct reference theory, this is strictly and literally false: since you already knew that Hesperus = Hesperus, it follows by (2) that you already knew that Hesperus = Phosphorus, and it is impossible to learn something that you already know. The descriptivist strategy provides a simple resolution of this prima facie contradiction: although the sentence is strictly false, there is something true that you asserted conversationally, namely, that you learned that the first heavenly body visible in the evening = the last heavenly body visible in the morning.

There is, however, an open-ended family of kindred prima facie contradictions which this approach, when generalized, cannot explain away. By way of illustration, suppose that you just read *Naming and Necessity* for the first time and found yourself persuaded at every step. You tell us:

- (5) I learned from *Naming and Necessity* that necessarily Hesperus = Phosphorus.

⁹ *Naming and Necessity* (Cambridge: Harvard, 1980), pp. 142–44.

¹⁰ Cf. chapter 8 in Scott Soames, *Beyond Rigidity* (New York: Oxford, 2002).

We have the intuition that you have told us something true.¹¹ Given direct reference theory, however, (5) cannot be strictly and literally true. As before, since (we may assume) you already knew that necessarily Hesperus = Hesperus, it follows by (2) that you already knew that necessarily Hesperus = Phosphorus; but it is impossible to learn something you already know. To avoid this contradiction, direct reference theorists must identify the true thing that you asserted conversationally. Applying the above descriptivist strategy, we arrive at the following:

I learned from *Naming and Necessity* that necessarily the first heavenly body visible in the evening = the last heavenly body visible in the morning.

This paraphrase, however, is plainly false: it is *contingent*, not necessary, that the first heavenly body visible in the evening = the last heavenly body visible in the morning. Since you cannot (in the factive sense) learn something that is false, the paraphrase does not meet even the minimal requirement of producing some *true* thing that is conversationally asserted.

One might try to avoid the modal failure of this paraphrase by fiddling with scope: for example, perhaps you learned from *Naming and Necessity* that, for some x and y , x = the first heavenly body visible in the evening and y = the last heavenly body visible in the morning and necessarily $x = y$. But this too fails, for we may suppose that before reading *Naming and Necessity* you already knew the principle that identical objects are necessarily identical (as opposed to the new principles Kripke defended¹²) and that, from this principle and your recently acquired astronomical knowledge, you had already inferred the information reported in this paraphrase. But there is surely something *further* you (reported having) learned from *Naming and Necessity*.

There are other maneuvers that the descriptivist might try, but each of them falls prey to analogues of the litany of problems afflicting the

¹¹ The dialectical requirement from note 7 is therefore in force.

¹² See pp. 3–4, *Naming and Necessity*. Kripke says, “We must distinguish three distinct theses: (i) that identical objects are necessarily identical [that is, $(\forall xy)(x = y \rightarrow \Box x = y)$]; (ii) that true identity statements between rigid designators are necessary; (iii) that identity statements between what we call ‘names’ in actual language are necessary. (i) and (ii) are (self-evident) theses of philosophical logic independent of natural language.... From (ii) all that strictly follows about so-called ‘names’ in natural language is that *either* they are not rigid *or* true identities between them are necessary” (p. 4). So plainly, amongst these theses, (iii), *not* (i), is the *new* semantical principle Kripke tells us he discovered. (I pick up this theme two paragraphs hence.) The point in the text is that, from (i) and what you learned from the pamphlet, you cannot infer the *new* thing you reported learning in uttering (5).

description theory of names itself—modal failure, failure of serious publicity, lack of phenomenological plausibility, and so on. For example, the modal failure of the above descriptive paraphrase can be avoided by inserting actuality operators. But, in this case, serious publicity would be jeopardized: for example, according to the resulting report, what you learned from Kripke could not possibly have been learned in any alternate world just like the actual world except for some insignificant difference outside your light cone.¹³ And so it goes, problem after problem. The pattern that seems to emerge is that the requirements a successful descriptive paraphrase must satisfy if it is to capture what you asserted with (5) and the requirements a successful description theory of the names 'Hesperus' and 'Phosphorus' must satisfy are effectively the same. Accordingly, if direct reference theorists *could* find a descriptive paraphrase for (5) that avoids the analogues of all the problems afflicting the description theory of the names 'Hesperus' and 'Phosphorus', evidently the constituent descriptions would provide the means of identifying the senses of 'Hesperus' and 'Phosphorus'. If this generalizes, then the description theory of names and the descriptivist paraphrase strategy stand or fall together.

There is a related difficulty that befalls the descriptivist strategy. When Kripke uttered 'Necessarily, Hesperus = Phosphorus', he took himself to be asserting some *new* thesis—something he learned from philosophical reflection (plus background knowledge). Presumably, what he asserted was the very thing he took himself to have learned. Moreover, what he asserted was presumably what you took him to have asserted, what you took yourself to have learned from him, and what you asserted that you had learned from him. All these were one and the same thing. Now since, according to the descriptivist paraphrase, what you asserted was really descriptive, it follows that what Kripke asserted when he uttered 'Necessarily, Hesperus = Phosphorus' was likewise descriptive. The same thing goes for the arguments Kripke gave on behalf of his assertion: since his assertion was

¹³ Cf. Soames, *Beyond Rigidity*, pp. 39–50. A similar problem befalls paraphrases that invoke other rigidifying devices (for example, "thick words"; see note 23). Likewise, it is not the case that what you learned was that 'Necessarily, Hesperus = Phosphorus' is true in English. For reasons, see my "The Origins of Modal Error," *Dialectica*, LVIII, 1 (January 2004): 11–42, on pp. 29–30. By way of illustration, suppose that you reported 'Marie learned that, necessarily, Hesperus = Phosphorus', that Marie learned what she did by reading an essay on Kripke's modal metaphysics written entirely in material mode, and that you have no knowledge of the language in which it was written or the language Marie speaks. Surely, what you reported Marie to have learned is not the indicated metalinguistic fact.

descriptive, the premises of his arguments for it must (on pain of accusing Kripke of a fallacy of equivocation) have been descriptive as well. But the resulting picture cannot be right: it goes against everything Kripke took himself to be arguing (and everyone else took him to be arguing) in *Naming and Necessity*.

This brings us to a simple but serious defect in all candidate descriptivist paraphrases (a defect they share with the descriptive senses of names proposed by Fregeans), namely, they are just not plausible. When one utters sentences of the form 'I learned that $A = B$ ' or 'I learned that $A \neq B$ ' intending to assert something conversationally, although there may sometimes be a plurality of asserted propositions, plainly there are at least some cases in which one *consciously* has in mind a single most salient thing (a "primary assertion") that one is *consciously* asserting. And among such cases, there are at least some in which the conversational assertion one is consciously aware of is nondescriptive. Think about it. For example, someone utters 'I learned that Cicero \neq Ovid' and is consciously aware of what he is asserting conversationally. Must he really be consciously asserting anything like the following: that he learned that the most famous Roman orator \neq the author of *Metamorphoses*? Surely not. In such a nonidentity, where there is no theoretical pressure to accept any such descriptive paraphrase, it is immediately evident that such paraphrases are at least sometimes unfaithful to the phenomenology. In such cases, what he consciously asserted was simply that he learned that Cicero \neq Ovid. This is our way of saying it. Once we have come this far, it then seems incredible that, if he goes on to utter 'I also learned that Cicero = Tully' (or better still: 'I learned that Cicero is not Ovid but, rather, Tully'), what he is consciously asserting should suddenly become descriptive. And, of course, such a view would be plainly unfaithful to the phenomenology.

A final problem with the descriptivist paraphrase strategy is that it seems to have lost track of the greater dialectical situation and, as a result, risks ending in a stalemate with traditional Fregean descriptivism.¹⁴ By way of illustration, suppose our direct reference theorists bite the bullet and choose to avoid the above modal failure by allowing rigidifiers into their paraphrase of (5) (as we discussed above). In

¹⁴ An analogous threat of stalemate befalls Kripke's own use of descriptivist paraphrases in connection with his attempt to deflate intuitions reported with 'It could have turned out that Hesperus is not Phosphorus'. Fortunately, his arguments do not require these paraphrases, since he also has a *nondescriptive* paraphrase strategy, which avoids stalemate. See my "Mental Properties," this JOURNAL, xci, 4 (April 1994): 185–208; and "The Origins of Modal Error," pp. 24–28.

this case, Fregean descriptivists would themselves be entitled to use rigidifiers to paraphrase (and thereby deflate) reports of various *prima facie* anti-Fregean intuitions. The following is one way they might try to do this: hold that names really express *nonrigid* descriptive content and then hold that names occurring in reports of *prima facie* anti-Fregean intuitions are *not* being used strictly and literally and that they are instead being used to express *rigidified* descriptive content.¹⁵ This is not to say that the resulting view would be plausible; it is not. The point is that the overall implausibility of this description theory of names supplemented with these descriptivist paraphrases is not obviously greater than the overall implausibility of direct reference theory supplemented with its descriptivist paraphrases. Let us not forget the implausibility of the immediate consequences of tenets I and II (for example, that you have known all along, and indeed knew *a priori*, that necessarily Hesperus = Phosphorus) and the phenomenological implausibility of the paraphrases introduced by direct reference theorists to mitigate this implausibility. If indeed the two theories do not differ in overall plausibility, stalemate results.

In view of the foregoing worries, the appeal of the prevailing three-place approach sketched in section II becomes clearer. For it seems to have a straightforward way of dealing with sentences like (5). Given a three-place treatment of learning, the true thing you conversationally asserted when you uttered (5) would be that you learned *under a mode of presentation similar to 'It is necessary that Hesperus = Phosphorus'* that it is necessary that Hesperus = Hesperus. Unlike the descriptivist paraphrase, this paraphrase does have claim to being true. This is just one of the many attractions of this view.

IV. REDUCING PROPOSITIONS TO ORDERED SETS

We come now to the second method for avoiding the three-place treatment of knowledge. Instead of treating 'know *a priori*' as a three-place predicative expression 'Know-*a priori*'³ (x, p, m), it is treated (at least in its relevant pragmatic uses) as two-place, but now the

¹⁵ To illustrate how this would go, consider the *prima facie* anti-Fregean intuition uncritically reported with 'It is possible for Hesperus to be bumped out of its orbit and no longer be visible from earth, in which case Hesperus would not be identical to the first heavenly body visible in the evening'. A Fregean who holds that 'Hesperus' is synonymous to the nonrigid description 'the first heavenly body visible in the evening' would, in the present dialectical setting, be free to identify what is asserted conversationally by uttering this sentence with (something like) this: it is possible that the *actual* first heavenly body visible in the evening be bumped out of its orbit and no longer be visible from earth, in which case the first heavenly body visible in the evening [in this possible situation] would not be identical to the *actual* first heavenly body visible in the evening.

satisfiers of the second place are ordered-pairs (sequences, labeled trees) consisting of a Russellian proposition p and a mode of presentation m of p : 'Know-a-priori'² ($x, \langle p, m \rangle$).¹⁶ There is an obvious intuitive objection to this proposal. When something is known a priori, that which is known must be a genuine bearer of truth and falsity (necessity and possibility, and so forth); that is, that which is known a priori must be a proposition. But, intuitively, an ordered pair cannot be true or false (and so forth). Advocates of the present approach have two replies to this problem.

First, they may treat the indicated occurrence of 'Know-a-priori'² as a piece of technical terminology (about "a priori schmowing," not a priori knowing). But then how are we to understand this technical term 'Know-a-priori'²? It seems that it must be explicated in terms of the three-place idiom 'Know-a-priori'³. But in this case our direct reference theorists are right back where they started.¹⁷

The second, and more serious, reply to this intuitive objection is that propositions really *are* ordered sets (sequences, labeled trees): philosophical analysis shows that propositions are *reducible* to such entities, and, in particular, some of them are reducible to the indicated ordered pairs $\langle p, m \rangle$. Consequently, these ordered pairs can be objects of knowledge, bearers of truth, and so forth.

There are various problems with this position. First, such reductions simply fail, for they are subject to fatal Benacerraf-style problems.¹⁸ For example, these direct reference theorists hold that, when we assert (1) in conversation, we pragmatically assert that it is not possible to bear the relation knowing-a-priori² to a certain proposition, which in turn is taken to be an ordered set. But which ordered set is it? Is it the ordered pair $\langle \langle \text{Hesperus}, =, \text{Hesperus} \rangle, \langle \text{Hesperus} = \text{Phosphorus} \rangle \rangle$? Or is it $\langle \langle \text{Hesperus} = \text{Phosphorus} \rangle, \langle \text{Hesperus}, =, \text{Hesperus} \rangle \rangle$? Or $\langle \langle \text{Hesperus}, \langle \text{Hesperus} \rangle \rangle, \langle =, \langle ' = ' \rangle \rangle, \langle \text{Hesperus}, \langle \text{Hesperus} \rangle \rangle \rangle$?

¹⁶ A more sophisticated, perhaps superior, alternative to $\langle p, m \rangle$ is the ordered pair $\langle p, \sigma \rangle$, where σ is the equivalence class of modes of presentation similar to m .

¹⁷ Another problem with this reply is that, in the context of two-place theories, the basic tenets of direct reference theory have the immediate consequence that, when we know that ϕ , we stand in a relation of knowing to the proposition that ϕ ; but in the case of "a priori schmowing" we stand in some technical relation to a mere ordered set. This observation also serves to block a reply based on the so-called "fallacy of incomplete analysis" (according to which my intuitive objection in the previous paragraph turns on wrongly mixing analyzed and unanalyzed terms). See my "Propositions," p. 5.

¹⁸ See my "A Solution to Frege's Puzzle," *Philosophical Perspectives*, vii (1993): 17–61. See also Joseph Moore, "Propositions, Numbers, and the Problem of Arbitrary Identification," *Synthese*, cxx, 2 (August 1999): 229–63; and Michael Jubien, "Propositions and Objects of Thought," *Philosophical Studies*, civ, 1 (May 2001): 47–62.

And on and on and on. No answer stands out as objectively correct: each is equally satisfactory, and equally unsatisfactory, as the others.¹⁹ A natural response to this problem is to abandon the idea that propositions are genuinely *reducible* to ordered sets (sequences, labeled trees) and to hold instead (as John Perry, Jon Barwise, and many others have) that these set-theoretical constructs serve only to *represent* irreducible, *sui generis* propositions.²⁰

Second, direct reference theorists (whether or not they are moved by this Benacerraf-style objection) cannot recognize propositions of the sort representable by (or reducible to) the indicated ordered pairs $\langle p, m \rangle$. For, as indicated earlier, direct reference theorists recognize only pure Russellian propositions. But, as the present proposal is intended, all Russellian propositions are representable by (or reducible to) either sequences of properties and objects $\langle F, a \rangle$ or sequences of n -ary relations and objects $\langle R^n, a_1, \dots, a_n \rangle$ or more complex sequences formed from such sequences with the aid of logical particles—for example, $\langle \text{negation}, \langle F, a \rangle \rangle$. But (where p is a Russellian proposition) the ordered pair $\langle p, m \rangle$ is obviously not formed in such a way, and so there is no Russellian proposition representable by (or reducible to) it. (Hereafter, I will simply use the idiom of representation.)

Indeed, there is good reason why direct reference theorists do not recognize the non-Russellian propositions that would be represented by $\langle p, m \rangle$. For if there were such propositions, then among them would be propositions suitable for a successful purely semantical solution to Frege's Puzzle—the alleged impossibility of which provides a central motivation for direct reference theory. (Or, more cautiously, once direct reference theorists entertain the contemplated propositions, their standard Kripkean arguments against the existence of a purely semantical solution to Frege's Puzzle would no longer be effective and so direct reference theory would simply lose its underlying motivation.) Among the indicated non-Russellian propositions would be propositions with three key features that enable them to function

¹⁹ The analogous point holds for the reductions of Russellian propositions explicitly endorsed by most direct reference theorists, for example, their reduction of the proposition literally expressed by 'Hesperus = Phosphorus'. Is this proposition the ordered triple $\langle H, =, H \rangle$ or $\langle =, H, H \rangle$ or $\langle H, H, = \rangle$? Or is it a labeled tree? Or a partial function f on positive integers such that $f(1) = \text{identity}$; $f(2) = H$; $f(3) = H$? Or what? There simply is no objectively correct answer.

²⁰ Of course, until one has a *theory* of this relation of representation (holding between these set-theoretical constructs and genuine propositions), this approach does not constitute a genuine *theory* of propositions; it is only an heuristic tool.

successfully as the senses of sentences.²¹ First, since mode m has *fine-grained* identity conditions, $\langle p, m \rangle$ inherits those fine-grained identity conditions and so, in turn, does the non-Russellian proposition represented by $\langle p, m \rangle$. Second, since mode m can be *nondescriptive* (for instance, m might be just 'Hesperus = Phosphorus' itself), the proposition represented by $\langle p, m \rangle$ can itself be nondescriptive, as Kripkeans require. Third, since the modal value of the Russellian proposition p would be inherited by the non-Russellian proposition represented by $\langle p, m \rangle$, the latter proposition would have the *modal value* (that is, necessity) required by Kripkeans as long as p has that modal value. So, for example, if m and m' are distinct nondescriptive modes of presentation, then the propositions represented by $\langle [\text{Hesperus} = \text{Hesperus}], m \rangle$ and $\langle [\text{Hesperus} = \text{Hesperus}], m' \rangle$ would be distinct, nondescriptive, and necessary, as Kripkeans require.

If there are indeed non-Russellian propositions with these features, one would have the makings of a purely semantical solution to Frege's Puzzle that is immune to the Kripkean critique. All that would be required is that there be a family of nondescriptive modes of presentation that respect the relevant intuitive distinctions. Presumably, there would be modes like this among the sorts of modes m direct reference theorists themselves invoke in their conversational pragmatics. For example, m might be an equivalence class of suitably related pairs of a sentence and a possible language (for example, $\langle \text{'Hesperus} = \text{Phosphorus'}$, English \rangle).²² Or m might be an equivalence class of suitably related "thick sentences"—that is, sequences of "thick words" (that is, words that have their referents essentially).²³ Obviously, there

²¹ This and the ensuing points about propositions can be generalized to individual concepts.

²² For a (rough) illustration, let us call S a "proto-synonymy" relation holding between sentence/language pairs iff S would always accord with pretheoretic intuition regarding questions of synonymy and nonsynonymy (indeterminacies allowed). Assume that S is an equivalence relation. (If it is not, replace S with some explicated variant that is.) S then determines equivalence classes m . If sentence/language pair $\langle s, L \rangle$ belongs to m , then let us stipulate that s "has proposition q as its sense" in L iff q is the proposition represented by $\langle p, m \rangle$ (where p is the relevant Russellian proposition). Clearly, there are relations S meeting these conditions. A Frege Puzzle arises when pretheoretic intuition judges a pair of true sentences ' $\alpha = \beta$ ' and ' $\alpha = \alpha$ ' to be different in meaning. But such differences in meaning—and, hence, Frege Puzzles themselves—could then be explained by the fact that ' $\alpha = \beta$ ' and ' $\alpha = \alpha$ ' have different senses (for every S , if there is more than one). Since, by design, the propositions that would be the senses of ' $\alpha = \beta$ ' and ' $\alpha = \alpha$ ' (where α and β are names or common nouns) are nondescriptive and necessary, this semantical solution circumvents the standard Kripkean arguments against traditional sense-based solutions.

²³ See my "A Solution to Frege's Puzzle"; and Kit Fine, "Essence and Modality," *Philosophical Perspectives*, VIII (1994): 1–16.

are other candidates.²⁴ It is hard to see what could rule out there being *some* such modes that do the job.

In any case, at this stage of the dialectic, if direct reference theorists are willing to countenance the general family of non-Russellian propositions under discussion, they will have exhausted their resources for arguing that there exist no modes *m* such that the associated pairs $\langle p, m \rangle$ represent the sort of non-Russellian propositions that could function as senses of sentences in a purely semantical solution to Frege's Puzzle. The upshot is this. Since direct reference theorists are committed to there being no purely semantical solution, they must outright reject the general family of propositions representable by (or reducible to) ordered pairs of the type under discussion.

V. PROPOSITIONS AND THE LOGIC OF IDENTITY

We have just seen that, in the context of direct reference theory, the two leading two-place treatments of the problematic conversational uses of 'knows', 'knows a priori', and so forth are not tenable. This brings us back to the prevailing three-place approach. There are a number of standard problems facing this approach. My present purpose, however, is to raise a very different sort of problem having to do with the logic of the identity relation.

Consider the negation of the original premise (2):

- (6) The proposition that Hesperus = Hesperus and the proposition that Hesperus = Phosphorus are *not* identical.

In symbols:

$$\neg([\text{H} = \text{H}] = [\text{H} = \text{P}]).$$

To most people (even most people who know the astronomical fact

²⁴ For example, if *p* is the Russellian proposition that direct reference theorists believe to be expressed by 'Hesperus = Phosphorus', then *m* might be a descriptive proposition of the sort Fregeans believe to be expressed by 'Hesperus = Phosphorus'. In fact, two-dimensional semantics is really just a variant of this idea, on which the "secondary intension" is the coarse-grained possible-worlds "proposition" *p'* necessarily equivalent to *p* and the "primary intension" is the coarse-grained possible-worlds "proposition" *m'* necessarily equivalent to *m*. Many two-dimensionalists believe that the proposition really expressed by 'Hesperus = Phosphorus' is the ordered pair $\langle p', m' \rangle$. But, because of Benacerraf-style problems, $\langle p', m' \rangle$ is not a genuine proposition but is at best a mere *representation* of a proposition. Moreover, as in note 20, until we have a theory of this relation of representation, this approach does not yield a genuine theory of propositions. Nor do two-dimensionalists come any closer to a genuine theory when they call the indicated primary and secondary intensions "aspects" of the proposition they take to be expressed by 'Hesperus = Phosphorus'—at least not until they provide a genuine theory of this relation of *being an aspect of* (on a par with set theory, property theory, λ -calculus, and the like).

and who are convinced by the Kripke-Putnam arguments against descriptive senses), (6) certainly *seems* true.²⁵ But tenet II implies that (6) is literally false. Indeed, (6) is just (an instance of) the very premise that direct reference theorists explicitly reject in their “solution” to the logical version of Frege’s Puzzle. So once again we have a *prima facie* contradiction that needs to be explained away. Given their pragmatic explanatory strategy in section II, our direct reference theorists must by parity hold that, when one assertively utters (6) in conversation, what one is saying conversationally would be true. But what could this true thing be? To maintain uniformity with their treatment of ‘knows a priori’ as three-place (thereby providing an additional place for a mode of presentation), our direct reference theorists are evidently committed to adopting an analogous pragmatic treatment of relevant conversational uses of the identity predicate itself. In particular, they evidently would have to treat such conversational uses of the identity predicate as *four-place* (thus providing two additional places for modes of presentation—one for a mode similar to ‘Hesperus = Hesperus’ and the other for a mode similar to ‘Hesperus = Phosphorus’). Thus, the conversational use of the unembedded occurrence of the identity predicate in (6) would have to be represented as a four-place predicate:

$$\neg(=^4([H = H], m, [H = P], m')).$$

Enough is enough!

Virtually all philosophers hold that identity is binary. (Even those few philosophers who entertain a Geach-style relative identity theory hold that identity is ternary, not quaternary.) In any case, I know of no direct reference theorists willing to reject the standard binary treatment of identity. Indeed, the binary character of identity looms large in the original Kripkean arguments—the very arguments that initially motivated the abandonment of Fregean semantics in favor of direct reference theories. Just try winning Kripke’s conclusions while at the same time treating identity as four-place. For me, however, the most important problem with this proposal is this: I just cannot understand what this pragmatic paraphrase is supposed to be saying. The reason will become clear in a moment.

But I am getting ahead of myself: direct reference theorists still have some escape routes for avoiding the four-place treatment of

²⁵ One may, if one wishes, restate (6) using naked ‘that’-clauses without the technical term ‘proposition’. Moreover, since we have the intuition that (6) is true, the dialectical requirement from note 7 is in force.

identity. First, they could try to avoid the entire issue by just refusing to generalize from their three-place treatment of the propositional attitudes to this four-place treatment of identity. But this proposal is unacceptably ad hoc: the attitude puzzle associated with (1), for example, and the puzzle associated with (6) certainly *seem* to be of the same general sort. To put the point the other way round, if it is unacceptable to multiply argument places in the case of identity, is it not equally unacceptable to multiply argument places in the case of the attitudes? And, of course, by making this ad hoc move, the account of intuitional error loses its systematicity. Finally, if direct reference theorists reject the four-place treatment of identity, they still owe us an explanation of the *prima facie* contradiction between (6) and tenet II.²⁶

Another way of trying to avoid the four-place treatment of the indicated conversational uses of the identity predicate is effectively the same as the unsatisfactory two-place, ordered-pair treatment of conversational uses of 'know a priori' (considered in the previous section). On this treatment, the true thing asserted conversationally by uttering (6) would be something like:

The ordered pairs <the proposition that Hesperus = Hesperus, 'Hesperus = Hesperus'> and <the proposition that Hesperus = Hesperus, 'Hesperus = Phosphorus'> are not identical.

Besides being extremely implausible that any such thing is what we would really mean by uttering (6) in conversation, the proposal cannot be extended to various slightly more complex cases, for example,

²⁶ For example, on analogy with the proposal discussed in note 7, they might claim that what we intuit when (6) is asserted conversationally is really false (and really does lead to contradiction) but then try to account for the source of our error by treating intuiting as a ternary relation. But this proposal fails to meet the dialectical requirement discussed in that note. Moreover, to maintain uniformity with the emended account at the close of that note (which does meet this requirement), the present account would need to be emended analogously, so that the truth of what we understand and intuit is ensured by incorporating modes of presentation into its propositional content (one similar to 'Hesperus = Hesperus' and the other similar to 'Hesperus = Phosphorus'), thus leading us right back to the four-place treatment of identity.

Another, quite different proposal is to hold that the true thing we hear when (6) is asserted is simply that the relevant modes of presentation (versus the propositions) are nonidentical, for example, that 'Hesperus = Hesperus' \neq 'Hesperus = Phosphorus'. But (6) is tantamount to the central premise of the logical version of Frege's Puzzle—the very thing being denied in tenet II. It is extremely implausible that parties to the debate are hearing assertions of (6) as not even being about propositions. Are we to believe that we have forgotten the very thing under debate? And, in any case, this proposal falters over slightly more complex examples like that discussed in the following paragraph.

utterances of 'The proposition that Hesperus = Hesperus can be known a priori and therefore is not identical to the proposition that Hesperus = Phosphorus, which cannot be known a priori'. The present proposal yields the following pragmatic paraphrase:

The ordered pair <the proposition that Hesperus = Hesperus, 'Hesperus = Hesperus'> can be known a priori and therefore is not identical to the ordered pair <the proposition that Hesperus = Hesperus, 'Hesperus = Phosphorus'>, which cannot be known a priori.

But this sort of paraphrase is ruled out for exactly the chain of reasons detailed in the previous section.²⁷

As no doubt has occurred to many readers, there is a final, quite natural alternative way of dealing with (6). Suppose that m is a mode of presentation similar to 'Hesperus = Hesperus' and m' , a mode of presentation similar to 'Hesperus = Phosphorus'. The proposal is to retain the canonical binary treatment of identity and to treat 'Under modes of presentation m and m' respectively' as a kind of *sentential operator* $U_{m,m'}$ that operates on ordinary identity and nonidentity sentences.²⁸ In conversation, then, when one utters (6), which is literally false according to tenet II, the true thing one would be saying conversationally would be:

Under modes of presentation m and m' , respectively, the proposition that Hesperus = Hesperus and the proposition that Hesperus = Phosphorus are not identical.

(For simplicity, I am suppressing the similarity clauses.) Without a multiplication of argument places, this is neatly represented thus:

$$U_{m,m'}([H = H] \neq [H = P]).$$

A further advantage of this treatment is that it naturally generalizes to the original puzzles involving the attitudes. The resulting treatment restores uniformity and avoids the multi-place treatment of all these familiar predicates. I therefore take this to be the most promising treatment of identity for propositions available to direct reference theorists.

²⁷ Instead of trying to avoid the four-place treatment of identity by means of the ordered-pair approach of section IV, a direct reference theorist could, of course, try to make use of the descriptivist approach of section III. But, predictably, this runs into many of the problems we have already considered, for instance, failure to handle various sentences that deal simultaneously with learning, necessity, and identity.

²⁸ See, for example, Graeme Forbes, "Objectual Attitudes," *Linguistics and Philosophy*, XXIII, 2, (April 2000): 141–83, for the idea of treating 'under mode of presentation m' as an operator.

This proposal, however, suffers from a fatal flaw, which is the point I have been building up to all along. According to direct reference theory, the proposition that Hesperus = Hesperus and the proposition that Hesperus = Phosphorus are identical. Let ' q ' name this proposition. Then it follows (by substituting ' q ' for ' $[H = H]$ ' and ' $[H = P]$ ') that the last indented formula is necessarily equivalent to: ' $U_{m,m'}(q \neq q)$ '.²⁹ In English: 'Under modes of presentation m and m' , respectively, q is not identical to q '. But I really do not understand what this is supposed to be saying! The kernel sentence ' $q \neq q$ ' is as patently false as a sentence can get. What relevant action of the adverbial operator $U_{m,m'}$ could possibly turn ' $q \neq q$ ' into a true sentence: $U_{m,m'}(q \neq q)$? I just do not get it.

To see my problem, consider the analogous but simpler sentence: 'Under modes of presentation d and t , respectively, Hesperus is not identical to Hesperus' (' d ' for dawn; ' t ' for twilight). In symbols: $U_{d,t}(H \neq H)$. But what could this sentence mean? When I try to understand it, I find myself saying the following; I just do not know how else to take it:

Hesperus at dawn is not identical to Hesperus at twilight.

where 'Hesperus at dawn' and 'Hesperus at twilight' are functioning as *singular terms*.³⁰ But this is exactly the sort of *false* (or nonsensical) thing one or two of our undergraduates may be counted on to say upon first hearing Frege's Puzzle! Direct reference theorists (and pretty much everyone else) certainly take this to be false, not true. Given this, 'Under modes of presentation d and t , respectively, Hesperus is not identical to Hesperus' must itself be false. But this sentence and our original, more complex sentence 'Under modes of presentation m and m' , respectively, q is not identical to q ' have exactly the same form: $U_{u,v}(\alpha \neq \alpha)$. Accordingly, the more complex sentence must be saying:

$$q\text{-under-}m \neq q\text{-under-}m'$$

where ' $q\text{-under-}m$ ' and ' $q\text{-under-}m'$ ' are singular terms. Therefore, this sentence must be false for the very same reason that 'Hesperus at twilight is not identical to Hesperus at dawn' is false. In turn, 'Under

²⁹ Does $U_{m,m'}$ generate a referentially opaque context? No, not according to direct reference theories: the reference of a proper name occurring in an otherwise opaque context is, strictly and literally, never opaque but instead is always "direct." Bear in mind that the sentences being used as paraphrases are supposed to be taken strictly and literally.

³⁰ Or, alternatively, 'Hesperus presented at dawn is not identical to Hesperus presented at twilight'.

modes of presentation m and m' , respectively, q is not identical to q' must likewise be false. This sentence, however, was supposed to literally express the *true* thing conversationally asserted by uttering (6). So this account of (6)—and of identity statements for propositions generally—leads to inconsistency.³¹

For many philosophers, this would be the end of the matter. Suppose, however, that some direct reference theorists try to swallow this treatment. How, in that case, would they propose to understand the singular terms ' q -under- m ' and ' q -under- m' '? One answer, of course, is that these terms refer to ordered pairs $\langle q, m \rangle$ and $\langle q, m' \rangle$, respectively. But the resulting view would not be relevantly different from the ordered-pair proposal considered in the previous section and, therefore, must be rejected for the same reasons. A second answer is that they refer to objects in a highly novel metaphysical category, namely, "*qua*-objects." *Qua*-objects are discussed by Aristotle and various medieval philosophers (according to whom, this road *qua* from Athens to Thebes \neq this road *qua* from Thebes to Athens). On this view (stated roughly), x -*qua*- F is much like the ordinary object x except that the property F , which is not an essential property of x simpliciter, is an essential property of x -*qua*- F . The idea, then, would be that there are, not only *qua*-*particulars*, but also *qua*-*propositions*: for example, the proposition that Hesperus = Hesperus *qua* m and the distinct proposition that Hesperus = Hesperus *qua* m' . But, if there truly exist *qua*-propositions, one would evidently be free to adopt the hypothesis that these *qua*-propositions are identical to the propositions denoted by the associated 'that'-clauses ('that Hesperus = Hesperus' and 'that Hesperus = Phosphorus', respectively). In this case, the proposition that Hesperus = Hesperus and the proposition that Hesperus = Phosphorus would not be identical, contradicting tenet II of direct reference theory.³² And, given the equivalence of tenets I and II (as seen in section I), tenet I would be false as well. The upshot is that, if to explain away the *prima facie* inconsistencies concerning identity that result from tenets I and II, direct reference

³¹ And, since $\neq (q, m, q, m')$ iff $U_{m,m'}(q \neq q)$, the four-place treatment of identity fails for the very same reason.

³² By the same token, if there were really such things as *qua*-propositions, it would seem that there should also be corresponding *qua*-concepts—for example, the concept *qua* m of being Hesperus and the concept *qua* m' of being Hesperus. If so, these *qua*-concepts could then be identified with the missing Fregean senses of the proper names 'Hesperus' and 'Phosphorus'. Of course, someone might propose to reduce *qua*-propositions and *qua*-concepts to ordered pairs. But, then, we would once again be right back to the proposal rejected in section IV.

theorists resort to qua-objects, tenets I and II—and so direct reference theory itself—would still come out false.

VI. CONCLUSION

In summary, the basic tenets of direct reference theory lead to straightforward *prima facie* contradictions, which direct reference theorists must explain away. They propose to do this by means of a kind of pragmatic explanation. We showed, however, that their usual two-place approaches to such explanations collapse, leading them to the prevailing three-place approach. But when this approach is generalized to deal with certain examples concerning the identity of propositions, an intractable dilemma results. Either they must treat identity as a four-place relation (contradicting what just about all philosophers, including direct reference theorists themselves, hold about identity). Or they must incorporate a view that was rejected in pretty much our very first lesson about identity—namely, that Hesperus at twilight is not identical to Hesperus at dawn. In desperation, some direct reference theorists might attempt to escape the dilemma by embracing an ontology of qua-objects, but this only opens up the possibility of a purely semantical solution to Frege's Puzzle and so is tantamount to abandoning direct reference theory. The overall conclusion, therefore, seems to be that the direct reference theory is forced into a treatment of identity statements that is inconsistent with basic principles concerning the logic of identity, which nearly all of us, including direct reference theorists, take as starting points.

GEORGE BEALER

University of Texas at Austin