A Multi-Model Modal Theory of I-Semantics
Part 2: Identity and Attitudes

Christopher Tancredi
KEIO UNIVERSITY

1. Model Theory

Model-theoretic semantics provides the basis for formalizing the semantics of natural language. As its name implies, model theory is concerned with how to model meaning in natural language. A model will be sufficient for this purpose if it provides the tools needed for giving adequate semantic interpretations to all sentences uttered by all speakers of the language it models.

For the purposes of this paper, I take a model to be a tuple consisting of a set of individuals D, a set of worlds W, and an evaluation function F, i.e. <D,W,F>. The model is used to give an interpretation to expressions in a language L. The evaluation function F associates the basic expressions of L with extensions in each of the worlds in W – names with individuals, one-place predicates with sets of individuals, n-place predicates with n-tuples of individuals, etc. While it is picturesque to view the worlds of the model like we view the world we live in, filled with real people smelling real flowers and real stars exploding in real supernovae, for the purposes of model theory this is not a requirement. All that a world has to do in model theory is to make it possible to determine, together with F, the extensions of basic expressions in a language. While model theory is in many respects non-intuitive and one may justifiably shirk away from taking model theory to give the true meaning of language, it does provide a useful tool for exploring aspects of meaning related to truth and entailment and in this sense serves an indispensable purpose in developing a semantics of natural language. The sketch just given of model theory ignores several things that are required in a fully

adequate account, minimally including a set of times and also plausibly a set of events. These limitations reflect the focus of this paper, not substantive claims about what primitives a model must contain.

Standard approaches to model-theoretic semantics attempt to employ a single model in interpreting expressions of a language. This usually tacit decision I will argue makes it impossible to use model theory for what we most want to get out of our semantics, namely a way of understanding and evaluating utterances by different speakers. I will propose in this paper that abandoning the tacit single-model assumption in favor of a theory that can employ distinct models to model the interpretations assigned to sentences by different speakers makes it possible to adapt model theory to these purposes. In addition to giving us a way of modeling understanding and evaluation of other speakers, a multi-model approach to semantics also makes possible a straightforward analysis of the distinction between *de re* and *de dicto* interpretation.

2 Single-Model Theories

The core of a model is the set of individuals it contains, D. If D contains two individuals, say *a* and *b*, then it follows from the definition of a set that $a \neq b$. We may well have multiple names for these individuals, calling *a* A and calling *b* either B or C. If this is the case, "B is C" will be true, but this does not reflect any identity between distinct individuals. At the level of the individuals such a statement merely asserts that $b=b$. "A is B" in this case would be false, since it asserts that $a=b$, an impossibility if *a* and *b* are distinct elements of D. It is this property of models that I will show leads to problems in understanding other speakers. For all speakers are susceptible to confusion about identity, but a single model cannot show us what such confusion consists in.

Consider for illustration the situation of two distinct speakers, Lois Lane and Clark Kent, who have an irreconcilable difference in their ontologies. Where Lois sees two individuals *c* and *s*, Clark sees only one. It isn't that Clark isn't looking close enough. In a literal sense, he can see everything that Lois can. However, when Lois looks at Clark she sees a reporter, and when she looks at Superman she sees a superhero who she takes to be distinct from that reporter. Clark, on the

other hand, simply sees himself in two different costumes playing two different roles. Given this distinction between Lois and Clark, it is clear that in one sense what they will mean in uttering a sentence containing the name *Clark* or *Superman* will differ even if the sentences they employ are syntactically identical. At the same time, we will evaluate those sentences identically, taking both to be true or both to be false depending on how we interpret the sentences ourselves. An adequate theory of semantics should be able to describe this state of affairs. In particular, it should be able to give us a way of *understanding* the utterances of the two speakers as different as well as a way of *evaluating* the utterances identically.

Now suppose we wanted to model utterances by Lois and Clark in model theory using a single model. What would the domain of individuals in that case be? In particular, how many individuals would there have to be in the domain of individuals corresponding to the names *Clark* and *Superman*? We can imagine 3 possible answers to this question. A realist approach would say there is only one such individual, say a, since in fact Clark is Superman, Lois's beliefs notwithstanding.[1] A name-centric approach would say that there are two such individuals, c and s, one associated with each name. Finally, a mixed approach would allow there to be as many individuals in the domain as are needed to account for both Lois's and Clark's world views, in this case three: a, c and s. Every one of these approaches is problematic.

The realist approach is perhaps the most widely assumed in the philosophical literature, and is at the heart of theories of direct reference and rigid designation. On this approach, there is a unique fact about the reference of a name in a language, and if multiple names refer to the same individual then they are semantically synonymous. When Lois says "Clark isn't Superman", under the realist approach she makes a statement whose interpretation is of the form $a \neq a$. This gives us an immediate explanation for the intuition that Lois is wrong in what she says. More generally, the realist approach excels at explaining our evaluation of utterances by other people, at least on the surface, since it disregards the different intentions underlying the utterance and focuses merely on the realist interpretation of the utterance itself. However, the strength of the realist approach in evaluation is its weakness in understanding. In the present case, it runs headlong into Frege's

problem: how can Lois's claim that Clark isn't Superman be distinguished cognitively from the claim that Clark isn't Clark, a claim that Lois is not prepared to make? The realist approach assigns the same interpretation to both sentences, and in fact lacks the tools needed to distinguish them. That is, the realist approach fails miserably when it comes to trying to make sense of speakers' intentions. In particular it gives us no way of making charitable sense of Lois's utterance, i.e. of seeing it as anything other than an obvious and straightforward logical contradiction.

Frege of course was well aware of the above type of problem and proposed to account for our different attitudes to distinct but synonymous semantic truths by appealing to a distinction between sense and reference. Sense and reference are for Frege intimately connected, with the former determining the latter. On this view, evaluation could be made at the level of reference, and understanding at the level of sense. Whatever the advantages of this theory elsewhere, as Kripke (1980) has shown in eviscerating detail it fails miserably for accounting for the interpretation of names. The problem in a nutshell is that there is nothing that can serve as the sense of a name without thereby making obviously incorrect predictions about the meanings of sentences containing the names.

The problems with the realist model do not stop at understanding other speakers. More devastating is that such a model is divorced from speakers' intuition. For no speaker can justifiably claim to employ the realist model as a model of the language they speak. Misidentification of a single individual is enough to make the meaning we assign to an utterance of our own distinct from the meaning that the realist model assigns. Clearly a realist model cannot be used to model Lois's use of language since she fails to grasp a crucial fact about the domain of that model. It would be nothing less than the baldest act of hubris to claim immunity from such mistakes ourselves. And since there is no way of protecting against misidentification, we cannot justify taking our intuitions to tell us anything about the model. Even such universally recognized truths as that for any truth-evaluable statements *p* and *q*, the statement *p and q* entails *p* cannot be guaranteed to follow on a realist model. For we have no way of determining whether our intuitions about meaning are intuitions about the language as modeled

by the realist model, or even how the language we understand and use relates to the realist model. Without being able to determine this relation, native speaker judgments can count for nothing in probing the nature of the realist model, and hence under the realist view cannot be used to tell us anything about meaning in language.

Equally damning to the realist view is the observation that even if it could be claimed to give a correct picture of the semantics of some abstract E-language such as English, in order to understand people who attempt to use this language it will be necessary to set up an entirely separate model of the languages these people actually use, i.e. their I-languages in Chomsky's sense. This is so since without such a model we are at a complete loss for explaining how to characterize Lois's confusion. Within the presumed correct realist model for English, the names *Clark Kent* and *Superman* both refer to a single individual, namely *a*. This means that the value assignment function F assigns the individual *a* as value to these two names. Within such a model there is no way of explaining what Lois's confusion consists in. To say that she doesn't realize that Clark is Superman would be to say that she doesn't realize that $a=a$. Taking Lois's relation to the realist model into consideration doesn't do the trick either. From that perspective we could perhaps say that Lois doesn't know the value of F(Clark) or of F(Superman), or perhaps of either, or that she simply fails to grasp the model as a whole, but that would leave a mystery the fact that Lois does appear to understand a wide range of sentences involving these names and it would do nothing to help us understand what Lois IS grasping. Since Lois is guilty of no internal inconsistency, it must be that she has a way of representing things in which her statement that Clark isn't Superman has the form $x \neq y$, and not the form $x \neq x$. But this characterization of the facts requires modeling Lois's way of representing things with something other than the realist model. Thus use of a realist model thus comes with a requirement of employing something outside of that model to make sense of utterances by people who do not grasp or employ that model, and the realist model tells us nothing about what that something could be.

The name-centric approach to modeling natural language fares no better than the realist approach. Under the name-centric approach, the model used for

interpretation contains one individual for each name. This approach makes it possible to understand Lois very easily. When she says "Clark is not Superman", her utterance is interpreted in the model as saying $c \neq s$, a logical truth. However, this approach makes evaluation of Lois problematic. Intuitively we want to say that Lois is incorrect, but under this approach doing so commits us to denying that $c \neq s$, i.e. to accepting that $c=s$, which within the model is a necessary falsehood. We face a similar problem in understanding and evaluating Clark. When Clark says (to himself, of course, not to Lois) "Clark is Superman," he makes a statement whose interpretation is $c=s$. However, such a statement cannot be true within model theory for the simple reason that no two distinct members of a set can be identical. That is, if $c=s$, it follows that c and s are the same individual, which goes against the initial assumption that all names are associated with distinct individuals within the model. Salvaging this approach requires an explanation for how Clark can be confused in making such a statement as well as for why we take his confused statement with truth conditions that cannot be satisfied to be true and Lois's straightforwardly analyzable "Clark isn't Superman," which is a necessary truth under this view, to be false.

The mixed view is at least as problematic as the realist and the name-centric views. When Lois says that Clark isn't Superman on the mixed view, her statement is intended as meaning that $c \neq s$. When Clark claims that Clark is Superman, in contrast, his intended meaning is that $a=a$. (A variant of this view would have Clark's statement mean that $c=c$, dispensing with the need for a, though this variant does nothing to overcome the problems noted here.) Here we have a straightforward way of representing both of the statements that we want to. Lois's "Clark isn't Superman" comes out as true since within the mixed model c and s are in fact distinct individuals. Likewise, Clark's "Clark is Superman" comes out true since the individual a is – like all individuals – self-identical. The problem here is that intuitively we would take Lois and Clark to contradict one another by uttering their respective statements, but on the present model there is no way of accounting for that intuition. The situation depicted would be no different from one in which Lois and Clark are in fact talking about distinct individuals with homophonous names. Quite generally, the connection between a name and an

individual in the model is entirely unconstrained on this approach, as is the connection between individuals in the model and individuals in the real world. And if one real world individual can be represented in the model by any number of distinct model-individuals, then the model can hardly claim to be a model of the real world.

The name-centric and mixed theories have little if anything to recommend them to begin with, and so as straw men may appear impossibly implausible. The point in examining them, however, was not to argue against plausible but mistaken positions maintained in the literature but rather to exhaust the range of possibilities even potentially worth considering of explaining natural language semantics using a single model in a model theory. Any other assumptions one could make regarding the domain of individuals in the model of English are even less plausible than the three possibilities considered here, and will not be considered individually. The above considerations could be taken as an argument against using model theory to model natural language semantics. However, dispensing with model theory would be throwing out the baby with the bath water. I propose instead to expand the way in which model theory is put to use.

The main proposal of this paper is that understanding and evaluation have to be able to involve distinct models. A model used for understanding Lois's utterances must incorporate her own presumed world view, including her presumed ontology of individuals as well as her presumed beliefs about those individuals. Under such a model, if the presumptions are correct then any sincere assertion by Lois will be interpreted as true, and hence understanding what the model is like will shed light on what Lois's mental state is like. A model for evaluating Lois's utterances, in contrast, needs to be based on the evaluator's view of reality. I will start with the assumption that a realist model can serve for evaluation, though ultimately this assumption will be seen to be untenable. I will argue instead that what is needed for evaluation is a model of the evaluator's world view. The resulting multi-model theory will then qualify as a theory of I-semantics, i.e. a theory of the interpretation of language with respect to the I-languages of individuals. Among other things, this approach will account for the uniformity of Lois's evaluations of other speakers, something that a realist approach cannot give

us.

3. Multi-Model Theory

I take as a starting point the assumption that model theory is the best theory we have for giving an interpretation to sentences of natural language, i.e. for associating a given syntactic form with a semantic value with respect to a given state of affairs. The realist view can be seen as the claim that the given state of affairs is fixed as the state of a single world, the actual world, and that this state of affairs fixes the domain of existing individuals as those individuals in that world. (Fictitious individuals can be introduced into the domain for the model as well, though obviously not as existing in that world.)

Lois is of course in a bit of a pickle with respect to a realist model since her assumed ontology is inconsistent with that of the realist model.[2] The rules of interpretation for the realist model applied to sentences she is prepared to utter yield some straightforward contradictions in that model. Part of the problem is no doubt the fact that Lois doesn't have a proper understanding of the state of affairs of the world she inhabits. But then what does she take her sentences to mean? The only way of elucidating this notion is by appealing not to the actual state of affairs in the real world but rather to the state of affairs that Lois takes to hold in the real world. For the actual world comes up at least one individual short of what Lois needs to explain the meanings of her sentences as she sees them. It will not do to say that Lois thinks of the one man who is both Clark and Superman that he is not himself. Nor would it do to say that Lois identifies one real world individual as Clark and a separate one as Superman, since one of these would end up as a misidentification by her own lights. Once Clark is fixed as Clark, there is no non-Clark individual that Lois identifies as Superman. Taking Lois's individuals to be fictitious individuals doesn't help either. For if Lois is taken to identify a fictitious individual as Superman, no assertion she makes about that individual could ever possibly be true in the actual world. Every *Superman* utterance she makes would have to be relegated to the realm of fiction, her own opinions on the matter notwithstanding. No, the best we can do in order to understand Lois is to model her language based on her own assumed ontology.

Then and only then does it become possible to show what she means when she says "Clark isn't Superman".

Understanding Lois of course comes at a cost. If we model Lois's language using her own personal ontology, we are faced with the burden of having to figure out how to make sense of the idea that anything she says can be about the real world and not a mere description of her own state of beliefs. This burden, however, can be met by relating individuals across two distinct models, one used for understanding other speakers and another used for evaluating them. There are two basic approaches that one could take to meeting this burden. One is to assume that the individuals of Lois's model are in principle distinct from those of the evaluation model, requiring some kind of correspondence relation to be set up not only for the problematic Superman/Clark Kent cases but also for all run-of-the-mill cases as well. This would be equivalent to adapting the Lewis (1968) theory of counterparts so that it applies to individuals in domains of distinct models rather than (or as well as) to individuals across worlds. However, there is no conceptual or technical reason why distinct models should have to be based on non-intersecting sets of individuals. The second approach does not make this assumption but rather assumes that except for the individuals for whom there is direct evidence for incompatible ontologies across speakers, the domain of individuals for Lois's model is identical to that for the realist model. This second approach has the advantage of requiring less abstraction on the part of the researcher, making explanations easier to give and to follow, so I will adopt that approach here. It also allows us to take rather direct advantage of the many insights into reference and its interaction with modality that Kripke (1980) has given us, while making it possible at the same time to account for the different behavior of modality and propositional attitudes and for the puzzles that the latter give rise to, such as those in Kripke (1979).

When Lois says "Clark is not Superman," how are we to understand her statement when on the realist model her statement is a straightforward contradiction? Charity dictates that we not interpret her as knowingly uttering a contradiction, which means that we have no choice but to conclude that the model that models her language use is not the realist model. This does not give us

blanket justification for using any model whatsoever for interpreting Lois, however. The evidence that we have of her language is in the vast majority of cases consistent with its being modeled by the realist model. Indeed, it may well be that the only evidence there is for a divergence from the realist model lies in her statement that Clark is not Superman. This suggests that the strategy we should employ for understanding Lois is to attribute to her a model that differs from the realist model as minimally as possible in such a way that within that modified model her utterance is (i) non-contradictory, and (ii) true. Satisfying (i) means using a model based on a domain containing two individuals, *c* and *s*, where the realist model contains a single distinct individual, *a*. Assuming the rest of the domain to be unaffected by Lois's misunderstanding about Superman and taking D' to be the set of individuals common to the Lois model and the realist model, the domain D_L of a Lois model will then consist of D' ∪ {*c*,*s*}, and that of the realist model D_R will consist of D' ∪ {*a*}. Satisfying (ii) will require not interpreting the sentence Lois utters with respect to the actual world, since *c* and *s* are by hypothesis not in that world.

Assuming a distinction in domains of course affects the sets of worlds contained in the two models. In particular, no world containing *a* in the realist model will be contained in the Lois model, and similarly no world containing either *c* or *s* in the Lois model will be contained in the realist model. But this is not problematic for understanding Lois. To understand Lois is simply to construct a model that fits Lois's world view, and there is no reason why we should be constrained to use only individuals and worlds from the realist model in doing so.

Understanding Lois must be separated from evaluating what she says for truth or falsity. Within the model used to understand Lois, the statement "Clark is not Superman" is true, denoting the proposition that $c \neq s$. However, our intuition tells us that the sentence is in fact false. If we used the Lois model for evaluating as well as for understanding Lois's utterances, we would never be able to explain how a sentence that Lois takes to be true could be seen by us to be false. Of course there is no reason why we should use Lois's model for evaluation, and there is good reason for not doing so. After all, we clearly see that she is missing something that we all see, namely that Clark IS Superman. Evaluation should

instead consist of interpreting her utterance with respect to a non-Lois model such as the realist model. In that model, her utterance is interpreted as denoting the proposition that $a \neq a$, and this proposition is clearly false. Putting this together formally, to both understand an utterance of Lois's and to evaluate a sentence uttered by Lois with respect to a realist model, semantic interpretation has to proceed as follows, where ϕ_{Lois} is an utterance of a sentence ϕ by Lois, w is the actual world, M_L is a Lois model, M_R is a realist model, and for any subscript *i*, $M_i = \langle D_i, W_i, F \rangle$.

(1) i. ϕ_{Lois} is understood as a claim in w that there is an M_L for which $[\![\phi]\!]^{M_L} = 1$
ii. ϕ_{Lois} is evaluated in w as true iff $[\![\phi]\!]^{M_R} = 1$

3.1 Refinements and Extensions

The multi-model analysis of interpretation just presented accounts straightforwardly for two distinct but equally important aspects of interpretation: understanding and evaluation. The analysis has some peculiarities, however, which may not at first glance be obvious. Most important is that neither understanding nor evaluation involves interpretation with respect to a specified world parameter. This is because it is a necessity of the analysis that at least one of understanding and evaluation not be made relative to the world of utterance. The reason should by now be obvious. If the world contains only one person where Lois sees two, then the world of utterance is not in a Lois model. Conversely, if there are two individuals in the world of utterance where the model used for evaluation sees only one, then the world of utterance won't be in the evaluation model. (Calling that model a realist model would be a misnomer in that case.) Since Superman / Clark cannot be one individual and two at the same time, this means there is no way for the world of utterance to be a part of both the Lois model and the realist model. And while we could try to overlook that fact and force interpretation with respect to the utterance world anyway, doing so would make it impossible once again to understand Lois. For if interpreted with respect to the utterance world (assuming the realist model to give the best view of what this world is like and in the absence of any operators affecting the world parameter),

Lois's utterance will be false. Understanding Lois, however, requires figuring out what things would have to be like for her utterance to be true, and the world of utterance simply does not help us to answer that question. Given the necessity of interpreting sentences without reference to a world parameter either in understanding or in evaluating Lois (or potentially both), the simplest assumption is that interpretation of a sentence uttered by Lois is never done with respect to a world parameter.

Formally, the absence of a world parameter has the potential to wreak havoc on interpretation. Standard theories of intensionality, after all, take all predicates to be world sensitive.[3] If we are to maintain the standard model-theoretic semantics of intensionality, then, there must be a way to provide a world parameter in order to incorporate world sensitivity into the theory. While the need for a world parameter is clear, however, it is not at all clear that the world of utterance ever needs to fill this role. The world of utterance does play an important role in determining the model. Interpreting Lois requires attributing to Lois a model of her beliefs,[4] and it is in the real world that those beliefs are fixed and held. However, this does not make the real world a parameter of interpretation.

To supply the needed world parameter for interpretation of an utterance I appeal to an intensional operator. This operator could be an overt modal or evidential. In the absence of an overt operator, I assume a default covert evidential operator Op_{evid}. Thus a simple declarative statement ϕ will have the form $Op_{evid}\ \psi$. Like all intensional operators, Op_{evid} quantifies over the set of worlds in the model. However, it adds a restriction to this set: a requirement of entailing evidence.[5] $Op_{evid}\ \psi$ is true with respect to a model M_L iff for every world w in W_L, in every world w' accessible to w such that the evidence in w obtains in w', ψ is true.

A note is in order regarding the notion of evidence used here. The term is intended to differentiate what an individual takes herself to know from what she merely believes. It is taken to include all evidence for anything, not merely evidence for the proposition claimed. The underlying idea is that a speaker will not take something to be part of her knowledge without believing herself to have entailing evidence for it. Evidence cannot be identified in a world by itself, but

rather needs to be fixed as an independent property of a world. In Lois's case, what Lois considers to be her own knowledge will be tagged as evidence in all of her belief worlds.

To fix ideas, let us now consider how it is that we understand and evaluate an utterance of another speaker. Consider the following utterance by Lois.

(2) Op_{evid} Superman flies and Clark doesn't

In order to understand this statement, we analyze it as a claim that there is a model M_L consistent with Lois's beliefs in the actual world according to which

(3) $[\![Op_{evid} \text{ Superman flies and Clark doesn't }]\!]^{M_L} = 1$.

This will be true iff for every world $w \in W_L$, in every world w' accessible to w such that the evidence in w obtains in w',

(4) $[\![\text{Superman flies and Clark doesn't }]\!]^{M_L,w'} = 1$.

Applying standard model theoretic computation from here, (4) will be true iff $[\![\text{Superman}]\!]^{M_L,w'} \in F(\textit{flies})(w)$ and $[\![\text{Clark }]\!]^{M_L,w'} \notin F(\textit{flies})(w)$. This will only be possible if $[\![\text{Superman}]\!]^{M_L,w'} \neq [\![\text{Clark }]\!]^{M_L,w'}$. If we assume that Lois is not attempting to utter a clear contradiction, then understanding Lois requires that we take D_L to contain distinct individuals as the values of $[\![\text{Superman}]\!]^{M_L,w'}$ and $[\![\text{Clark }]\!]^{M_L,w'}$. This tells us something about Lois's beliefs – that she takes the names *Superman* and *Clark* to denote distinct individuals. The utterance is then understood as a claim by Lois that she has evidence entailing that one of those individuals, the one she associates with the name *Superman*, flies and that the other, the one associated with the name *Clark*, does not. This is a plausible characterization of what we understand about Lois when she makes such a claim.

We turn now to evaluation. To evaluate Lois's utterance, we need to ignore Lois's ontology and the rest of her model. Under the approach outlined above, our evaluation derives from interpreting her statement with respect to the realist

model. Since Clark is Superman in this model, interpreting Lois's statement with respect to the realist model yields a contradiction. The relevant steps in the interpretation are as follows. First, evaluate Lois's statement as true iff

(5) $[\![\text{Op}_{\text{evid}} \text{ Superman flies and Clark doesn't }]\!]^{M_R} = 1.$

This will be true iff for every world $w \in W_R$, in every world w' accessible to w such that the evidence in w obtains in w',

(6) $[\![\text{Superman flies and Clark doesn't }]\!]^{M_R,w'} = 1.$

The identification of evidence in this case cannot be done by focusing on the beliefs of an individual since the realist model does not model an individual's beliefs. In a truly realist model M_R, however, there is exactly one world in W_R, the actual world w_a. In a sense, then, we can take everything in that world as evidence, constraining reality to the point that only one world remains as possible. In the unique world of the realist model, while there is plenty of evidence directly entailing that Superman flies, namely Superman's flying, there is no evidence entailing that Clark does not. Indeed, given the identity of Clark and Superman, evidence entailing the first conjunct entails the negation of the second. Thus in every world w' in which the evidence in w_a obtains, i.e. in w_a itself, (6) is a contradiction. And this gives us a reason to take Lois's claim to be false.

The adoption of the realist model for evaluation gives us a way to evaluate Lois's claims absolutely. All we need is absolute knowledge of what the world is like and we can make use of this model. It's an unfortunate limitation of the human species however, myself included, that such knowledge can never be attained. This casts doubt on any analysis that makes essential appeal to such a model. While it does give us an absolute judgment, we can never know what model is the realist model and so we can never use it to pass human judgment on other speaker's statements. And yet pass judgment we do. Fortunately, nothing in the above explanation depends on the model of evaluation being the realist model. The desired result, namely that we evaluate Lois as having said something

false, will follow if her statement is evaluated with respect to any model in which Clark is Superman, including a hearer's model. For such a hearer's model M_H, we evaluate Lois's statement as true iff for every world $w \in W_H$, in every world w' accessible to w such that the evidence in w obtains in w',

(7) [[Superman flies and Clark doesn't]]$^{M_H,w'} = 1$.

Here what counts as evidence is what the hearer takes himself as knowing, exactly parallel to the case of Lois. If that evidence entails that Superman flies and Clark doesn't with respect to the hearer's model, then Lois's statement is evaluated as true.

An added advantage of using a hearer's model to evaluate Lois is that it makes it possible to make a three way distinction in our evaluation: true, false, and possible. Suppose that Lois claims that Superman eats spinach. I who have no knowledge of Superman's gustatory habits cannot judge Lois's claim to be true, but nor can I judge it to be false. I simply lack the means of judging in this case. If we take the evidential operator to yield truth only when evidence entails a proposition and falsity only when evidence entails its negation, then we leave open the possibility of either truth or falsity in the in-between cases. This aspect of the interpretation of the evidential operator will only be relevant for evaluation, since understanding Lois's assertion still requires figuring out how she models the world on the assumption that her assertion is true with respect to her own model. Using a realist model for evaluation, on the other hand, does not allow for suspension of judgment.

The use of distinct models for understanding and for evaluation gives us a solution to Frege's problem, i.e. the problem of how a statement "A is A" could be cognitively distinct from a statement "A is B" when both are true. To say that both are true is to say that they are true in a model used for evaluation. To say that they can be distinct cognitively is to say that there are distinct ranges of models that make the two statements true: all models for the first, but only some models for the second. This in turn makes it possible for us to account for the fact that Lois, despite having uttered a contradiction by our lights, cannot be accused of

having made a logical error. The contradiction shows up only in the model used to evaluate her statement, and not in the model of her understanding of the statement.

While evaluation gives us the means of assessing Lois's statements, it comes up short in answering questions about Lois. In particular, evaluation does not tell us who Lois's beliefs are about, since no connection has been made between the Lois model containing the individuals she takes there to be and the realist model containing the individuals presumed to actually exist. To say that Lois's beliefs are about the real world individual who is known both as Clark and as Superman because that is how we interpret her sentence would under this approach be no more than an unjustified conjecture. If we take her beliefs to be about a real world individual at all, then, our account will at the very least need to be supplemented with a way of connecting this individual with the individuals of Lois's model.

The connection between individuals in the Lois model and those in the realist model cannot be forged mechanically. It rather requires justification. In the present case, the fact that Lois will point to *a* and call him *Clark* while denying that he's Superman when he's in certain attire, and will point to the same *a* and call him *Superman* while denying that he's Clark in other attire gives us the justification for associating our *a* with Lois's *s* and *c*. With this as justification, we can take any statement that Lois makes about Clark or Superman to be about *a*, this despite the fact that Lois will not see it as such.

4. Propositional Attitudes

The need to connect individuals across models shows up equally clearly in propositional attitude statements, and the solution just offered for the simple assertion case can be readily extended to such cases as well. Indeed, explicit recognition of multiple models makes it possible to give a very simple account of propositional attitude statements. In this section I outline an analysis of such statements that treats propositional attitude predicates as unambiguous, and in which the *de re* / *de dicto* distinction is purely a matter of scope.

To avoid attributing a contradiction to Lois, *de dicto* ascriptions have to

involve introduction of a Lois model in the interpretation of

(8) Lois believes that Superman flies and Clark doesn't.

The easiest way to do this is to analyze the predicate *believes* as triggering interpretation of its complement with respect to a belief model of its subject. If the sentence as a whole is evaluated with respect to a hearer's model, this will give the following:

(9) [[Lois believes that Superman flies and Clark doesn't]]^{M_H} = 1 iff $\exists M_L$ M_L models Lois's beliefs and $\forall$ w $\in$ W_L, [[Superman flies and Clark doesn't]]^{M_L,w} = 1[6]

This interpretation is straightforward, involving only the two separate models M_L and M_H, the latter giving us a basis for evaluating the statement as a whole that the former giving us what is needed in order to understand Lois. Supplemented with the justification in taking Lois's *s* and *c* to correspond to the realist individual *a*, we can even infer from this who it is (in our terms) that Lois is claimed to have a belief about. Note that this approach to *de dicto* interpretation is similar to but distinct from that offered by Frege. Like Frege, I take interpretation of the embedded clause to differ from that of the matrix clause. Furthermore, the analysis given can be analyzed as involving a distinction between normal reference and indirect reference. This is as far as the parallels can be pushed, however. The identification of what counts as normal reference and indirect reference clearly differs on the two accounts. Under Frege's account a sentence's normal reference is a truth value and its indirect reference is a sense. Here, in contrast, both references are equivalent to propositions, i.e. sets of worlds, though to (potentially) distinct propositions.[7] The "normal" reference involves interpretation with respect to one's own model, while the "indirect" reference involves interpretation with respect to a model attributed to the subject.

De re interpretation differs from *de dicto* interpretation in that a *de re* attribution of belief does not suffice to tell us what Lois's belief looks like. A *de*

re belief report tells us only who the individual is in our terms about whom Lois is claimed to have a belief. The form that that belief takes is not thereby specified. To recover that information requires a process of inference based on a justified connection between individuals in the Lois model and in the hearer's model. In the absence of such a connection, an analysis of *de re* attitude statements cannot be given without adding complications elsewhere. This can be illustrated with the following sentence.

(10) Lois believes that Clark flies

While clearly false if *Clark* is taken *de dicto*, the sentence is true under a *de re* interpretation of that name, as the following paraphrase helps to make clear:

(11) Lois believes of Clark (a.k.a. Superman) that he flies.

The problem we face here, both with the *de re* interpretation and with its paraphrase, is that on the one hand we want to use the hearer's model to identify Clark as *a*, but given the interpretation of the predicate *believe* needed to generate *de dicto* interpretations we need to refer to a Lois model to model her beliefs, and this model does not contain *a*.

To see how this problem affects model theory, consider the scope theory of *de re* interpretation, according to which *de re* interpretation is taken to follow from assignment of matrix scope to an element in the embedded clause. For concreteness I will assume that such scope assignment in the present case derives from LF raising of *Clark*. Following basic assumptions from Heim and Kratzer (1998), the naïve assumption would be that this LF raising results in the following structure:

(12) [Clark 1 [Lois believes [that t_1 flies]]]

The bare index is semantically interpreted as a predicate abstractor and the indexed trace as a variable. This will yield the following as an intermediate stage in the

interpretation, where g is an arbitrary variable assignment function with respect to which the sentence as a whole is interpreted, and g[x/1] is exactly like g except at most in assigning x as value to syntactic variables with the index 1:

(13) $\lambda x.$ [[Lois believes that t_1 flies]]$^{M_H,g[x/1]}$ ([[Clark]]^{M_H})= 1

Using the interpretation of *believes* from before, and assuming that [[Clark]]$^{M_H} = a$ and that variable assignment functions get carried over to new models introduced by attitude predicates, we derive the following as a later step in calculating the truth conditions.

(14) $\exists M_L$: M_L models Lois's beliefs and $\forall w \in W_L$, [[t_1 flies]]$^{M_L,w,g[a/1]} = 1$

Interpretation of the last piece will require that the value assigned by g[a/1] to the trace t_1 be a member of the set of individuals that fly in w, i.e.

(15) $a \in$ F(flies)(w)

However, this can never be the case for the familiar reason that *a* is not an individual in any Lois belief model and hence will not occur in any Lois belief world w, meaning that it cannot possibly be in the extension of the predicate *flies* at such a world. This means that in precisely those cases where we would most want to make a *de re* attribution to Lois, the attribution we make will come out false every time. Clearly the analysis needs to be modified.

In assigning *a* as the interpretation of *Clark*, the above analysis captures the intuition that *de re* interpretation is about individuals assumed by the evaluator (i.e. the hearer) to be real, not (necessarily) about individuals that the attitude holder takes to be real. However, the simplistic analysis employed of the semantics of *believes* makes it impossible to use this individual in characterizing Lois's beliefs, leading to the undesirable conclusion that all *de re* statements about Clark / Superman attributed to Lois will be false. There are three assumptions of this analysis that could be plausibly challenged: (i) that *de re* interpretation derives

from scope assignment, (ii) that *de re* interpretation employs the same predicate *believe* that was used above for the *de dicto* interpretation, or (iii) that the introduction of a new model carries over variable assignment functions unmodified. There are certainly ways of abandoning (i) that will still make possible a *de re* interpretation. However, no such approach will work unless it incorporates a way of associating a single expression with two distinct interpretations, one a member of D_H and another a member of D_L. Furthermore, if (ii) is left untouched, such a change will still require making adjustments to variable assignment functions to account for quantification into attitudes, i.e. it will still require us to abandon (iii). Thus abandoning (i) by itself is not sufficient. Unless it is found that abandoning (iii) alone is also insufficient, (i) should then be retained as the simplest analysis of the *de re* / *de dicto* distinction. Abandoning (ii) is technically possible as well, and is in fact a well-explored idea going back at least as far as Quine. Indeed the very expression "x believes *of y* that p" would appear on the surface to involve a version of *believes* that takes 3 arguments, contrasting with the 2 argument taking pure *de dicto* version used earlier. However, even if we admit an ambiguity in the predicate *believe*, we still have to face the problem of how to analyze mixed cases with one expression taken *de re* and another taken *de dicto* within the same embedded clause, and these cases are still going to require interpreting variables inside a proposition interpreted with respect to the subject's model and yet bound by an expression interpreted with respect to either a speaker's or a hearer's model. Unless such variable binding is built directly into the *de re* predicate, it is difficult to see how abandoning (ii) will help us with the problem, but building such variable binding into *de re* predicates is equivalent to abandoning (iii). These considerations make it clear that abandoning (iii) is at the very least necessary. I will show below that a proper understanding of variable binding into a *de dicto* environment is also sufficient to account for *de re* interpretation.

What remains is to characterize the change in the variable assignment function that is needed to deal with *de re* attributions. I will break this into a purely formal part and an intuitive part. A variable assignment function g I take to be constrained to assign to variables values within the model that is used for interpretation at the point at which g is introduced. If g is introduced at the matrix

level, then, the values it assigns to variables must be taken from the model used to interpret the matrix clause. For understanding a speaker, this will be the speaker's model, while for evaluating a speaker it will be the hearer's model. For illustrative purposes I will assume that these two models have identical domains so that there is no need at present to modify g for understanding or evaluation. In order to interpret variables inside a proposition attributed to Lois, g needs to be modified so that the values it assigns to variables occur in the domain of M_L. Formally, this requires changing g into a (potentially) distinct function g' such that for any n for which g is defined, g'(n) is a member of D_L.

I will employ a function G for making the required change in the variable assignment function, where both the domain and the range of G are the set of all possible variable assignment functions. The case we are interested in is the case in which the range of a given variable assignment function g is $D_{S/H}$ (the domain of individuals in the speaker's / hearer's model) and in which $D_{S/H}$ is not identical to D_L (the domain of individuals in a Lois model). In the case of those indices n for which g(n) is a member of $D_{S/H} - (D_{S/H} \cap D_L)$, we want G to alter the assignment so that [G(g)](n) is in D_L. Any such function G that converts variable assignment functions that are into one domain D_A into functions into another domain D_B I will call an **assignment conversion function**. In our *de re* interpretation of a belief statement, then, in order to be able to interpret the trace of the *de re* expression within the belief model of the attitude holder it is necessary to apply an appropriate assignment conversion function G to the variable assignment function g[a/1]. The relevant function that makes the statement true will be one for which G(g[a/1])(1) = *s*.

This gives us all the tools we need to be able to interpret attitude statements formally. We need only incorporate an assignment conversion function G into the semantics of attitude predicates. For the predicate *believe* we want something along the following lines, where D_u is the domain of expressions.

(16) $[\![\text{ believe }]\!]^{M,g} = \lambda p \in D_u \; \lambda x \in D_e \; . \; \exists M_x, G$: M_x models x's beliefs, G is a contextually justified assignment conversion function, and $\forall \, w \in W_x$, $[\![p]\!]^{M_x, G(g), w} = 1$[8]

With this as our semantics for *believe*, *de re* interpretation is straightforward. To illustrate, consider our earlier example. The interpretation of *believe* given above leads to the following derivation of truth conditions.

(17) a. $[\![(10)]\!]^{M,g} = 1$ iff $\exists M_L, G : M_L$ models Lois's beliefs, G is a contextually justified assignment conversion function, and $\forall w \in W_L$, $[\![t_1$ flies $]\!]^{M_L, G(g[a/1]), w} = 1$

b. $[\![t_1$ flies $]\!]^{M_L, G(g[a/1]), w} = 1$ iff $G(g[a/1])(1) \in F(\text{flies})(w)$

c. $[\![(10)]\!]^{M,g} = 1$ iff $\exists M_L, G : M_L$ models Lois's beliefs, G is a contextually justified assignment conversion function, and $\forall w \in W_L$, $G(g[a/1])(1) \in F(\text{flies})(w)$

Note that we cannot be satisfied with the mere existence of an assignment conversion function that allows us to interpret the complement of the attitude verb as true. Values of G such that $G(g[a/1])(1)$ is a bird or a plane make $[\![t_1$ flies $]\!]^{M_x, G(g[a/1])} = 1$ true, but those values are clearly irrelevant. The assignment conversion function that we choose thus must be justified. This justification brings a non-formal element into the analysis. In the present case, the fact that Lois will point to *a* and call him *Clark* while denying that he's Superman in certain attire, and will point to the same *a* and call him *Superman* while denying that he's Clark in other attire gives us the justification for associating our *a* with Lois's *s* and *c*. This in turn justifies converting variable assignment functions that map indices to *a* into otherwise identical variable assignment functions mapping those same indices to *c* or *s*. The justification is essential to generating plausible truth conditions. Without it nearly all belief statements would come out true.[9]

5. Advantages of the Multi-Model Approach

One of the decided advantages of the multi-model approach to semantics taken here over a single-model realist approach is that the present approach makes it possible not only to say what understanding Lois consists in but also what Lois's understanding of Clark consists in as well. By hypothesis, Lois is confused. She

sees Clark and Superman as two different individuals when in reality (according to the hearer) they are one. Clark, of course, is well aware of the fact that he is Superman. Now suppose that the time comes when Clark is tired of living a life of deceit and decides to come clean with Lois. He tells her flat out, "Clark is Superman. Superman is Clark".[10] He offers her no proof, however, and so Lois remains unconvinced. If Lois were constrained to using the realist model to model her understanding and evaluation of Clark's statement, then she would be at a loss. In this case, understanding could be adequately modeled, but evaluation would be left out. (This is, of course, the opposite from the situation of Clark's interpreting Lois's statement that Clark isn't Superman.) Using the multi-model approach, in contrast, makes it possible once again to model both Lois's understanding and her evaluation of Clark. Her understanding comes to finding a model M_C that would make Clark's statement true, hence one in which there is one individual, say *a*, who is known both as *Clark* and as *Superman*. Her evaluation of his statement, in contrast, is done with respect to her own model, and with respect to that model the claim is a necessary falsehood: $c=s$ cannot be true given that *c* and *s* are distinct individuals.

The analysis just given also extends to *de re* / *de dicto* contrasts. Lois is as capable of making a *de re* statement as is Clark. When her boss asks her to tell him what she knows about Superman, she tells him:

(18) Clark believes that Superman is Clark

Her intention is that *Superman* be taken *de re*. Unfortunately, the same statement could be analyzed *de dicto* and would be true under such an analysis, and it could still be taken to be about Superman once one adds the assumption that the Superman of Clark's model, i.e. *a*, corresponds to the Superman of Lois's, namely *s*. Fortunately, Lois was obliging enough to clarify her intention by following up with:

(19) Clark believes of Superman that he is Clark

The interpretation of this statement is straightforward. Understanding the statement consists of finding a Lois model in which the statement is true. Knowing what we do about Lois and Clark, such a Lois model will contain two individuals, *s* and *c*, where Clark sees one, *a*. We understand Lois as from her own perspective saying of *s* that Clark's correspondent to *s*, namely *a*, is identical to Clark's Clark. The assignment conversion function G chosen in this aspect of the interpretation will be one consistent with Lois's view of how individuals in her model connect in their various occurrences with those of Clark's model. We can only guess at (or rather stipulate) the value of this function, of course, but we have at least some good evidence to go on for making an educated guess. Evaluating Lois's statement consists in interpreting it with respect to a hearer's model, M_H. With respect to that model, Lois's statement will be true just in case Clark's correspondent to (the hearer's) $[\![\text{Superman}]\!]^{M_H}$ is identical to Clark's Clark. In this aspect of the interpretation we select a value for G that represents the hearer's view of how individuals in his model connect to individuals in Clark's model in their various occurrences.

Since evaluation depends on the hearer's model, it is possible for things to go wrong – or right – because of confusion on the hearer's part. If the hearer suffers the same confusion as Lois, the two of them will be happy as clams talking with one another, doubly confident in their (mis-)understanding of the world through their mutual reinforcement. The hearer will understand Lois as being right on target, and will evaluate her statement as true. If the hearer does not share Lois's confusion but rather shares the clear vision of (this thin slice of) reality that Clark enjoys, his understanding of Lois will show him that she is confused, this despite the fact that if he knows about Clark's beliefs he will evaluate her statement as true.

6. Conclusions

In this paper I have argued that a single model model-theoretic semantic analysis cannot give an adequate interpretation to identity statements. I proposed a modification of model theory that employs multiple models. In the simplest case, one model is used for understanding a person's utterance and a second model

is used to evaluate that utterance. It was then shown that this same analysis can be extended to give a natural account of propositional attitude statements in which the *de re* / *de dicto* distinction can be analyzed as a scope distinction without involving any ambiguity in the predicate. The analysis assigns the predicates an interpretation that directly captures *de dicto* readings, and accounts for *de re* interpretation by binding into the propositional complement of the attitude predicate. This binding was seen to require making a modification to variable assignment functions, with the exact modification made depending on context.

The analysis developed here gives a new solution to Frege's problem, i.e. the problem of how a statement of the form "A is A" can differ cognitively from one of the form "A is B" when the latter is true. The solution offered here is that the difference stems from the models with respect to which the statements are true. The former statement is true on all models. The latter, in contrast, is only true in models in which "A" and "B" denote the same referent. This means that ignorance about what model is correct can result in ignorance about the truth value of the latter statement, but cannot result in ignorance about the truth value of the former statement. Ignorance about a model should not be equated with contingency, of course. Within any given model, the truth value of "A is B" is not a contingent matter – it is the same in all worlds of the model, and hence whatever truth value it has in a world of that model it has necessarily. We know then of any identity statement that it has its truth value necessarily in any model, even though we may not know which truth value it is that it should have, a point made originally by Kripke (1980).

The analysis developed in this paper forms the core of a semantics for an I-language, or what I call I-semantics. It shows how to account for our core judgments of sentences uttered by others on the assumption that the model they presuppose may well differ from our own. Making the analysis work depends on differentiating between understanding and evaluation, and using different models for the two. Only by doing so can we analyze statements of other speakers as internally consistent and yet contradictory at the same time. In the relevant cases, consistency of a statement is with respect to a model that is taken to model the speaker's beliefs, while contradiction is derived by interpreting the same statement

with respect to a model of the hearer's beliefs. While the core parts of the analysis are in place, the task remains of applying it to a broader range of phenomena than have been addressed in this paper and integrating it into a more adequate semantics. For reasons of time and space, I put such applications off to a later date.

Notes

* I would like to thank Brendan Wilson and Chris Barker for insightful discussions on many of the topics written on here. I would also like to thank audiences at the Tokyo Conference on Psycholinguistics and at Gunma Prefectural Women's University for helpful comments.

[1] The objections to this first approach will carry over to what could be called a privileged perspective approach, where for example the interpreter's perspective is taken as definitive and the interpreter takes Clark and Superman to be one individual. A privileged perspective approach could be used for the name-centric approach as well with the interpreter sharing Lois's view of the non-identity of Clark and Superman. For our purposes, the differences between such an approach and those given in the text are inconsequential. They would share with the approaches examined in the text the assumption that interpretation involves only one model, and it is this assumption that I am challenging.

[2] I am taking the realist model in question here to be one based on a single world of the Superman type, i.e. a world in which the Superman story is true. The need to be able to talk as if within a story is already something that poses problems for a strict realist view in which the world used as the basis for the model is required to be the real one. This need can thus be seen as further motivation for adopting a multi-model approach to semantics. For the purposes of the present argument, however, it will suffice to assume that Lois and Clark are real world individuals, with Clark a flying superhero also going by the name Superman, thus avoiding the question of how to deal with truth in fiction.

[3] This is true even on an extensional-first model theory like that of Ben-Avi and Winter (2007), since in that theory, otherwise extensional predicates become

intensionalized in environments in which an intension is called for.

[4] A model models a person's beliefs if the individuals that comprise the domain of the model are those individuals believed to exist (actually and fictionally) and the worlds of the model are those worlds whose individuals are limited to those in the domain of the model and which are consistent with the beliefs of the speaker. Independent motivation for limiting the worlds of such a model to belief worlds comes from an analysis of asymmetries in the interpretation of doxastic and non-doxastic modals. The former have an obligatory wide-scope property that the latter lack. Making belief worlds part of the model used to interpret utterances makes it possible to explain this property of doxastic modals in a principled fashion by taking a modal base to simply be the set of worlds in the model. Analyzing the modal base as determined independently, in contrast, makes the special scope behavior of doxastics something that can only be stipulated. See Tancredi (2007) for details.

[5] See Williamson (2002) chapter 11 for an argument that merely very strong evidence is not sufficient. As he argues in detail, strong evidence can probabilistically lead one to a proposition as conclusion in as close to 100% of the cases as one would wish, but if all evidence one has for a proposition falls short of entailing that proposition we are not entitled to assert the proposition.

[6] I suppress the effects of the covert evidential operator in the matrix clause here and below for ease of illustration. Both Lois and her beliefs are identified with respect to the hearer's model M_H.

[7] In attributions of belief by a speaker to herself, the two propositions will be identical, though this cannot be guaranteed in the general case of attributions to others.

[8] Strictly speaking, the lambda expressions should be restricted to taking values from within the model M with respect to which *believe* is interpreted. I leave this restriction implicit here and throughout as it is orthogonal to the problem at hand. Since p is an expression and not a proposition, it is fair to assume that it will occur in all models with respect to which it is to be interpreted.

[9] We could do as well by introducing G in the semantics of *believe* as a free

variable whose value is to be pragmatically recovered from the context, and eliminating the existential quantification over G. This would yield the following as the interpretation of *believe*.

$[\![\text{believe}]\!]^{M,g} = \lambda p \in D_u \, \lambda x \in D_e . \; \exists M_x : M_x$ models x's beliefs and $\forall w \in W_x$, $[\![p]\!]^{M_x, G(g)} = 1$

For our purposes the two interpretations are interchangeable. I use the one in the main text primarily for expository purposes, since it makes the role of justification explicit.

[10] He is much more likely to say "I am superman", but interpreting such a statement requires introducing contextual parameters separate from models as in Kaplan's (1989) analysis, adding a degree of complexity to the explanation without thereby making anything any clearer. I will thus stick with a stilted Clark.

References

Ben-Avi and Winter (2007), A Modular Approach to Intensionality. ms., Technion, NIAS.

Frege, G. (1948) On Sense and Reference. *The Philosophical Review, 57, No. 3,* 209-230.

Heim, I. and A. Kratzer (1998) *Semantics in Generative Grammar.* Oxford: Blackwell.

Kaplan (1989) Demonstratives. H. Wettstein, J. Almog and J. Perry (eds.) *Themes from Kaplan* (pp. 481-564). Oxford University Press.

Kripke, S. (1979) A Puzzle about Belief. A. Margalit (ed.) *Meaning and Use* (pp. 239-283). Dordrecht, D. Reidel.

Kripke, S. (1980) *Naming and Necessity.* Cambridge: Harvard University Press.

Lewis, D. (1968) Counterpart Theory and Quantified Modal Logic. *Journal of. Philosophy, 65*, 113-26.

Quine (1956) Quantifiers and Propositional Attitudes. *Journal of Philosophy, 53*, 177-187.

Tancredi (2007) A Multi-Model Modal Theory of I-Semantics: Part I: Modals. ms., Keio University.

Williamson, T. (2002) *Knowledge and its Limits*. Oxford: Oxford University Press.