INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A-

BELIEF, NAMES AND MODES OF PRESENTATION:
A FIRST-ORDER LOGIC FORMALIZATION

by

RUILI YE

A dissertation submitted to the Graduate Faculty in Computer Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1999

UMI Number: 9946232

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

This manuscript has been read and accepted for the Graduate Faculty in Computer Science in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

6/2/99 Melvin Fitting Melvin C. Fitting
Date Chair of Examining Committee

Prof. Stanley Habib 6/2/99 Stanley Habib
Date Executive Officer

Richard Mendelsohn

Lawrence S. Moss

Stathis Zachos

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

BELIEF, NAMES AND MODES OF PRESENTATION:
A FIRST-ORDER LOGIC FORMALIZATION

by

Ruili Ye

Advisor: Professor Melvin C. Fitting

One issue which any natural language processing by the computers must face is the contexts that involve propositional attitudes such as belief. This is made more complicated when an object has more than one name, or the reference of a name does not exist. It is thus desirable to have a formalized, rigorous logic dealing with this issue, which may serve as the theoretical basis for future implementation. The present work is such a formalization. The semantical basis of the logic **FMP** is the notion of modes of presentation, introduced by Gottlob Frege in "On Sense and Reference". The language of **FMP** includes two belief operators, representing *de re* and *de dicto* beliefs, respectively, one of which is similar to the λ abstract used by Fitting and Medelsohn in *First Order Modal Logic.* We show that the axiom system of **FMP** is both sound and complete with respect to a natural semantics. We then apply it to some problems in the philosophy of languages, such as Kripke's puzzle about belief. This is a step toward building a "rational" robot that is able to reason about belief and naming as humans usually do.

Acknowledgement

My chief debt in writing this dissertation is to Prof. Melvin Fitting. Through the clarity of his style, Prof. Fitting not only taught the subject of logic, but also presented an excellent example of a veritable logician. As a constant source of knowledge and inspiration for ideas, Prof. Fitting's help throughout my thesis writing cannot be overstated: without his direct advice and guidance, this thesis would have been a nonexistent thing. I owe him more than I can possibly repay. It will also not be forgotten how he "stuck it out" to keep the appointments and normal teaching schedule, when he was attacked by a serious illness that would have kept others in complete bed rest for weeks. It is my special pride to have an advisor who is at once a scholar with a deserving name and a man of a most candid and upright character and genuine goodness, a man of justice, a person worthy of trust. Not a shred of hypocrisy, snobbery or nasty worldly tricks exists in him. Prof. Fitting is a true gentleman.

My initial attention to the subject of logic was purely the consequence of a sentence said to me by Prof. Michael Anshel: I am deeply indebted to his judgement, but for which things might have been quite different.

I wish to express my hearty thanks to Prof. Richard Mendelsohn, who provided many detailed and valuable comments on the philosophical part of the thesis, and to Prof. Larry Moss, who sacrificed much time going through the tedious proofs, and provided sincere and important comments and suggestions.

I'd also like to thank Prof. Rohit Parikh for introducing Frege's Puzzle to me and supplying numerous references, and for proposing to formalize modes of presentation (albeit by a dynamic logic approach). He also read the very early drafts of this formalization.

I am indebted to Dr. Gary Ostertag for his recommendation of several readings, in particular, Stephen Neale's *Descriptions*, one of the best contemporary works I have read in its area.

To Prof. Stanley Habib, I wish to express my great appreciation for his assistance as the Executive Officer, and for making the Graduate Teaching Fellowship available to me.

But grant that those can conquer, these can cheat;
'Tis phrase absurd to call a villain great:
Who wickedly is wise, or madly brave,
Is but the more a fool, the more a knave.
Who noble ends by noble means obtains,
Or failing, smiles in exile or in chains,
Like good Aurelius let him reign, or bleed
Like Socrates, that Man is great indeed. ...
A wit's a feather, and a chief a rod;
An honest man's the noblest work of God.
Fame but from death a villain's name can save,
As Justice tears his body from the grave,
When what t'oblivion better were resign'd
Is hung on high, to poison half mankind.

---Alexander Pope: *Essay on Man*

Contents

Introduction

In the computer processing of natural languages, it is inevitable to encounter contexts that involve such propositional attitudes as belief and belief reports. In such contexts, complications arise when an object has more than one name, or when a name does not denote an existent object. To study natural language processing and artificial intelligence in general, it is therefore of great importance to inquire into the meaning of belief and names, and examine the nature of belief report. These subjects have been studied in the philosophy of language for a long period of time, and are still hotly debated today; various theories have been developed by philosophers, among whom were some of the greatest in modern history, yet each theory has its advantages and defects. It would be of great interest to see some aspects of some major philosophical theories formalized into a rigorous logic system; for then, on the one hand, these theories themselves can be tested from a different point of view for their advantages and defects, and, on the other hand, such a formalization may some day serve as a theoretical basis for implementation for the computer. The present work is an attempt at such a formalization.

Chapter 1

Philosophical Background

1.1 A Historical Survey

1.1.1 Frege's Theory of Sense and Reference

In [13] and [14], Gottlob Frege investigated the nature of equality. In discussing the equality of content in Section 8 of [13], he noted that, unlike conditionality and negation, equality relates to names, not to contents. The symbol for equality of content, placed between two names, signifies that the two names have the same content. Consequently, every symbol has a double meaning with the introduction of a symbol for equality of content: it may stand for its content, or it may stand for the symbol itself. He then explains, with an example in geometry, why this does not imply that two symbols are not needed for the same content and why a symbol for equality of content is not an unnecessary invention, and concludes that a symbol for equality of content is needed because of the fact that the same content can be fully determined in different ways; "and *that*, in a particular case, *the same* content actually is given by *two ways of determining it*, is the content of a *judgment*." It is necessary, therefore, before making the judgment, to have two different names for the thing determined, which correspond to the two different ways of determination; it is also necessary, therefore, to have a symbol for equality of content to express the judgment, that is, to join the two names together. He continues: "It is clear from this that different names for the same content are not always just a trivial matter of formulation; if they go along with different ways of determining the content, they are relevant to the essential nature of the case." He introduces the notation

$$\vdash (A \equiv B)$$

to mean: "*the symbol A and the symbol B have the same conceptual content, so that A can always be replaced by B and conversely.*"

In [14], Frege begins by asking whether equality is a relation, and if it is, whether it is "a relation between objects, or between names or signs of objects". He found both views unsatisfactory: If it is a relation between objects, then the statement $a = b$ asserts the equality between what the names a and b designate, *i.e.*, that the object is identical to itself, therefore the statement would have the same cognitive value as the statement $a = a$, which is certainly not the case, since the statement $a = b$ is often very informative while $a = a$ never is; if, on the other hand, equality is a relation between names of objects, as he suggested in [13], then the statement $a = b$ would be a statement about the arbitrarily given names a and b, which would be absurd. What makes $a = b$ nontrivial, Frege observes, seems to be that it expresses the fact that the names a and b designate the same thing, *i.e.*, they are co-referential, and the fact that such a statement contains actual knowledge is due to the fact that the names a and b are not just arbitrary names, but each contains a *mode of presentation* of what is designated. Here Frege begins his distinction between two notions associated with a name: the *sense* and the *reference* of names (or signs). A name expresses its sense, designates its reference. To the name there corresponds a definite sense and to that in turn a definite reference, while to a given reference there need not belong only one name. It is in the *sense* where the *mode of presentation* is contained. For example, 'evening star' and 'morning star' have the same reference, the planet Venus, but they have different senses, containing different modes of presentation of the planet. Frege admits that there are cases where a name has sense but it's doubtful whether there is any reference corresponding to it, such as 'the celestial body most distant from the Earth'; thus, 'in grasping a sense, one is not certainly assured of a reference.' Some may say that we cannot talk of a name as if it refers to an object if we are not sure that it does have any reference; to this Frege replies that when we use a name in our speech, we presuppose a reference, even though we may be mistaken.

The problem raised thereby is commonly referred to as "Frege's Puzzle".

1.1.2 Russell's Theory of Descriptions

Frege's theory of sense and reference answers the question why it is worthwhile to make assertions about identity. However, Bertrand Russell, in his 1905 article "On Denoting", noted that in cases where a name does not have both a sense and a reference (or a meaning and a denotation), such a theory would confront difficulties. For example, in the sentence '*The present King of France is bald*', 'the present King of France' has a meaning but no denotation, and so the sentence seems to lack a truth value. This bothered Russell, to whom every sentence should have a truth value—if not true, false. Therefore Russell undertook to find a way to eliminate the problem. He suggested that we should not take the denotation to be what is concerned in propositions containing denoting phrases. He developed a theory of denoting based on the principle "that denoting phrases never have any meaning in themselves, but that every proposition in whose verbal expression they occur has a meaning." The theory regards a denoting phrase as essentially *part* of a sentence, and does not have any significance on its own account. Russell drew a distinction of scope. A description or denoting phrase has a 'primary' occurrence when the proposition in which it occurs results from substituting the description for 'x' in some propositional function ϕx; a description has a 'secondary' occurrence when the result of substituting the description for x in ϕx gives only *part* of the proposition concerned. A proposition in which a denoting phrase has primary occurrence is supposed to have first of all asserted the existence of that which is denoted; consequently, every proposition in which a description which describes nothing has a primary occurrence is false.

Once a logical theory is constructed, it needs to be evaluated, and one way to evaluate it is to test its capacity for dealing with puzzles. Having this view in his mind, Russell showed how his theory is able to solve certain puzzles which any theory of denoting ought to be able to solve.

Since to Russell descriptions belong with denoting phrases, and ordinary proper names are viewed as disguised definite descriptions, his Theory of Descriptions provided for him his theory of proper names.

A common ground shared by the theories of Frege and Russell is that ordinary proper names are treated as descriptional. However, there are significant differences between the two. To Frege, any proposition in which an object is referred to does not have the object itself as a component, but it refers to the object by virtue of its sense; only the sense, but not the object, enters into the proposition. But to Russell, all the constituents of a proposition are entities with which we have immediate acquaintance; when there is anything with which we do not have immediate acquaintance but only definition by denoting phrases, then the propositions in which the thing is introduced by means of a denoting phrase do not contain this thing as a constituent, but contain instead the constituents expressed by the several words of the denoting phrase.

Russell's Theory of Descriptions has been given a very clear exposition and forceful defense by Neale in [34], who also extended the theory.

The purpose of Russell's Theory of Descriptions, says Neale, is to make available a class of propositions to serve as the meanings of (utterances of) sentences of the form 'the F is G', whether or not anything answers to 'the F'. To explicate the Theory of Descriptions, Neale distinguishes two classes of propositions according to whether or not the proposition in question depends upon the existence of some particular individual. An *object-dependent* (or *singular*) proposition is a proposition about a particular individual that would simply not be available to be entertained or expressed if that individual failed to exist; in general, if 'b' is a genuine referring expression, then for any monadic predicate '—is G', the proposition expressed by an utterance of 'b is G' is object-dependent. A *descriptive* proposition is an object-*in*dependent proposition. Such a proposition can be expressed by an utterance of a sentence containing a description whether or not anything satisfies the description; and even if the description is satisfied, the proposition expressed can be grasped without knowledge of who or what satisfies it. The main claim of Russell's Theory of Descriptions which is defended by Neale, is that a phrase of the form 'the F' is not a genuine referring expression, and the proposition expressed by an utterance of 'the F is G' is object-*independent.* Because it treats descriptions as complex existential quantifiers rather than referring expressions, the Theory of Descriptions gave Russell the power to explain, without positing a realm of nonexistent objects, the fact that perfectly determinate thoughts may be expressed by utterances of such sentences like

> Ralph thinks the largest prime number is greater than 10^{29}
> John has never seen the present King of France
> The King of France does not exist
> Ponce de Leon searched for the Fountain of Youth

despite the fact that they contain nondenoting descriptions.

While advocating the Theory of Descriptions, Neale is one of those (including Kripke) who do not accept Russell's treatment of proper names. He noted that "endorsing the Theory of Descriptions *qua* theory of *descriptions* does *not* commit one to a treatment of proper names as disguised descriptions."

1.1.3 Kripke on Names and Necessity

In [24], Kripke gave an analysis of the notion of necessity and the nature of naming. He drew distinctions among three notions: necessity, *a priority* and analyticity. He argued that the tradition to identify the three notions—necessity, *a priority* and analyticity—was a mistake, and that it is wrong to regard ordinary proper names as disguised definite descriptions, as Frege and Russell did. Kripke's views on these subjects can also be found in [23]. To Kripke, names are always *rigid designators*; *i.e.*, a name designates the same object (if any) in all possible worlds. Even if one thinks of the reference of a name as being given by a description or a cluster of descriptions, there are two alternative theories. According to one theory, the name is regarded as *synonymous* with the description which replaces it, *i.e.*, the name *means the same* as that description or cluster of descriptions, thus the name will not necessarily designate the same object in all possible worlds, and this will make the theory of names into a theory of meaning. Alternatively, the description may be used to determine a rigid reference, in which case the referent fixed by the description will be the same in all possible worlds, and names will be rigid designators: what one has, then, is a theory of reference. This 'modern substitute' for Frege and Russell's theories that though a name is not a disguised description, it either abbreviates, or anyway its reference is determined by, some cluster of descriptions is wrong as well, Kripke argued this by showing that it fails in one way or another to correctly characterize names. He also criticized Frege for using the term "sense" in two different senses, taking the sense of a designator to be its meaning as well as the way its reference is determined, thus supposing both are given by definite descriptions.

The way names are treated—as rigid or nonrigid designators—is crucial in determining the meaning, and therefore, the truth or falsehood, of statements involving names and necessity. If we think of names as always being rigid designators, then the identity $Hesperus = Phosphorus$ states a necessary truth, since Hesperus and Phosphorus designate the same object, the planet Venus, and in all possible worlds this object (and any object) is identical to itself. The natural consequence of such a view is that all identities are necessary: there are no contigent identities.

Kripke makes a careful distinction between the notion of *necessity* and the notion of *a priority*. The concept of necessity, he states, is a notion of metaphysics, while the concept of a priority is an epistemological notion; the two cannot be identified with each other. That the standard meter in Paris is one meter long is known *a priori*, but it is not a necessary truth, because the length of that stick at a certain moment t_0 is used to *fix the reference* of the measuring unit 'meter', rather than to *give the meaning* of 'meter'. Through this distinction, Kripke departs from the traditional views before him which had treated necessity and *a priority* as two equivalent notions.

1.1.4 Information Content and Frege's Theory of Sense and Reference

In [49, 50], Salmon seeks to understand Frege's theory of sense and reference in terms of the *information content* of sentences. [49] is a defense of the *naïve theory* (which holds that the reference of a name occurs as a constituent of the propositions in which the name appears). Based on the view that the fundamental semantic role of a declarative sentence is to encode a piece of information, and that the central and fundamental semantic value of a declarative sentence is its information content, the piece of information encoded, he defends the thesis that "the thoughts we have and the propositions we assert, believe, *etc.*, when formulatable using ordinary proper names, are always Russellian 'singular propositions', in which the only thing contributed by a name's occurrence is the named individual, and ... the attributions of thought, assertion, belief, *etc.*, we make using proper names do nothing more or less semantically than ascribe thought, assertion, belief, *etc.*, toward just such propositions". To solve Frege's puzzle, it's important to realize that it is not a puzzle about identity, since Frege's puzzle can arise in connection with any predicate, not just identity (*e.g.*, 'Hesperus is a planet if Phosphorus is' vs. 'Phosphorus is a planet if Phosphorus is', and 'Shakespeare wrote *Timon of Athens*' vs. 'The author of *Timon of Athens* wrote *Timon of Athens*'). By Salmon's account, identity statements are only a special case of the general problem, which is a problem concerning pieces of information; more specifically, it concerns the analysis of the sort of information that is semantically contained in declarative sentences, the feature of sentences that account for their informativeness or uninformativeness. What Frege called "thought", in Salmon's view, is the same thing as the information content of a sentence. "The information value of any compound expression ... is made up of the information values ... of the information-valued components of the compound."

In [61], Taschek criticized such an approach to Frege's puzzle. He regards the tradition followed by Salmon and many others as a mistake from its *starting point* by viewing that the fundamental semantic role of a declarative sentence is to encode a piece of information, and therefore approaching Frege's theory of sense as if what led him to distinguish sense from reference were an issue about differences in *information content*. Instead, Taschek seeks to understand Frege's distinction between sense and reference and the significance of such a distinction by way of *understanding* Frege in the whole. "For Frege, we are obliged to distinguish sense from *Bedeutung* in order to do justice to certain differences between sentences which, by his lights, it is the business of *logic* to account for." Frege's real puzzle is, Taschek claims, how can two sentences that require in order to be true that the same circumstances obtain nevertheless differ *in the logically relevant way* that Frege supposed that such sentences as $a = a$ and $a = b$ differ? To understand Frege's puzzle, it is of utmost importance to recognize that Frege took differences in cognitive potential to manifest logically relevant differences in the contents of the relevant sentences. Were Frege to allow himself to talk about information content at all, says Taschek, he would grant that the pair of sentences have the same informaiton content, and the issue cannot be whether or not two sentences differing in cognitive potential differ in their information content, but rather, it should be that sentences of the puzzling sort differ in a logically relevant way *despite* the fact that they encode the same information. The fundamental semantic role of a declarative sentence is to encode or express a content which is essentially to be individuated in terms of its logical potential, and the real challenge posed by Frege's puzzle is to account for the way in which the logical properties of the contents of our sentences can outstrip their referential truth conditions in this way.

1.1.5 Puzzles about Belief

From Frege's puzzle there emerged other puzzles in the contexts of propositional attitudes such as 'knowing', 'believing'. In [25], Kripke presented a puzzle about names and belief. Pierre, a normal French speaker who has never been to London, concludes his impression of London from what he has heard about the city by saying

(1) *'Londres est jolie,'*

from which we may conclude:

(2) Pierre believes that London is pretty.

Later, having been to London and lived in an unattractive part of it, Pierre, without realizing the city which he now lives in and knows as 'London' is the same city he has called *'Londres'*, concludes in English, without realizing the fact that 'London' and *'Londres'* refer to the same city:

(3) 'London is not pretty,'

from which we may conclude:

(4) Pierre believes that London is not pretty.

The puzzle lies in that, on one hand, since Pierre is a normal, rational person, he should never believe that London is pretty and that London is not pretty at the same time; on the other hand, as (1) and (3) present no contradiction to Pierre, he does not withdraw his assent from (1) while uttering (3), and (2) and (4) are reasonable conclusions from Pierre's sincere utterances and so are justified.

In [60], Taschek argues that the above is in fact not a puzzle at all from Frege's point of view. To the Fregean, two sentences may differ in their cognitive significance solely for the exchange of co-referential names despite the fact that they are the same everywhere else. When two sentences differ in their cognitive significance, it is possible for a competent and reasonable speaker to believe one while disbelieving the other, for this is precisely what is implied by difference in cognitive significance. Admitting this, one sees that Kripke's puzzle is no puzzle at all for the Fregean, because when asking *us* whether or not Pierre believes that London is pretty, Kripke must suggest what cognitive significance, or sense, are associated with 'London' by us—the sense which (we take to be the same sense as) Pierre associates to 'Londres' or the sense (we take to be the same sense as) Pierre associates to 'London'; if the former is the case, the answer is yes; if the latter is the case, the answer is no; and there is a third case, *i.e.* , if we are unsure about what senses Pierre associates to the two names, we will be unable to answer the question. In any case, Taschek says, "this is no puzzle."

He set out to solve the puzzle in [62] with the aid of notions he called *local* and *global* logical structures. There are, says Taschek, two elements which together provide the basis for a natural and compelling semantic explanation of the deeply felt intuition of most speakers that the substitution of coreferential expressions in the content-clauses of belief ascriptions is not guaranteed to be truth preserving. The first is the so-called Logic Requirement (LR): "to guarantee that the content-sentences of our (oblique or *de dicto*) belief reports adequately and systematically reflect the logically relevant features of what our subjects believe, we standardly require that the content-sentences themselves possess, on the occasions of their use, corresponding logical properties." The second element, which was already argued in [61], is that "coreferential substitution typically results in a sentence with different logical properties from those of the original—surely a semantic and not a merely pragmatic difference." He claims that the substitutional opacity of belief ascriptions is a direct consequence of the fact that they are subject to LR given that coreferential substitution generally alters the logical properties of a sentence. "Insofar as we require ... that the logical properties of their content-sentences accurately mirror the logical properties of the beliefs being reported, then, since coreferential substitution cannot be guaranteed to preserve the logically relevant features of a sentence, it will come as no surprise that coreferential substitution within these content-sentences is not generally truth preserving." He proceeds to represent the logical structures of sentences by logical schematas, which explicitly show the difference between the logical structures of two sentences, one of which results from the other solely by a substitution of the coreferential expression, and thus show why the two sentences differ in logical potential. Two sentences are said to share the same local logical structure just in case there is no schema of which one is an instance and the other is not. Two sentences S and S' have the same global logical structure just in case for all logical schemata Σ and Σ' and any sentence S^*, S and S^* are corresponding instances of Σ and Σ' (resp.) just in case S' and S^* are also corresponding instances of Σ and Σ' (resp.). Taschek observes that the failure to preserve global logical structure is the cause of the failure to preserve the logical features of a sentence in coreferential substitution, which, in turn, leads to the general invalidity of coreferential substitution in the content-clauses of attitude ascriptions. In solving the puzzle about Pierre's belief about London, "a sensitive appeal to considerations of global logical structure will provide us with precisely what we need," and to do so, we must adopt a more holistic and context sensitive approach to evaluating the correctness of these belief ascriptions. It should be required that the content-sentence used to report Pierre's belief expressed by asserting (1) have a global logical structure which absolves it from logical inconsistency with the content-sentence of (4). Thus, once we have already committed ourselves to asserting (4), it would be more proper to report Pierre's belief in asserting (1) by something like the following sentence than by (2):

(5) Pierre believes that *Londres* is pretty,

"explaining, if need be, that Pierre doesn't know that *Londres* and London are one and the same.

In [12], Forbes also does not take the puzzle to be a puzzle about belief. Forbes supported Frege's idea of associating a sense with a name,while departing from the historical Frege by holding that we are not *completely* specifying the proposition the believer is said to believe in such sentences like

Lois believes that Superman can fly and Clark Kent can't.

Forbes claims that the puzzles presented by Kripke and others are not puzzles about belief, but quandaries about belief ascription. To use another of Kripke's examples, that about Peter and Paderewski, should we infer both that Peter believes that Paderewski had musical talent and that Peter believes that Paderewski had no musical talent? If the interpretation is uncertain, it is argued, one is under no obligation to answer such question as "Does Peter, or does he not, believe that Paderewski had musical talent?" This is why Fregean approaches provide the most plausible explanation to why substitution fails.

1.1.6 Representations of Belief

Hintikka pioneered giving strict logical formulations of the notions knowledge and belief. In [18] he systematically studied the criteria of consistency of statements containing knowing and believing, expressed by the operators K_a and B_a, where a is the knower or believer, and anticipated the idea of formulating knowledge and belief in terms of the possible world notion. The work has been a base for the developement of logics of knowledge and belief.

Traditionally, as in [18], believing is regarded as a two-place relation, holding between the believer and what is believed; among those based on this ground, Schiffer ([53, 54, 55]), in an attempt to solve puzzles similar to the example Kripke gave about Pierre and London, and borrowing from Frege the notion *modes of presentation*, distinguished two different ways to ascribe a sentence like the following:

(6) Ralph believes that Fido is a dog.

The "*Russellian* view" takes the references of 'Fido' and 'dog' to be Fido (the dog) and doghood, respectively, and the reference of 'that Fido is a dog' to be the Russellian proposition $\langle Fido, doghood\rangle$, and represents (6) in the form

$$B(Ralph, \langle Fido, doghood\rangle).$$

The "*Fregean* view" takes the references of 'Fido' and 'dog' to be modes of presentation of Fido and doghood, and represents (6) in the form

$$B(Ralph, \langle m_f, m_d\rangle),$$

where m_f and m_d are modes of presentation of Fido and doghood, repectively. Schiffer contended that the "Russellian view" will not be able to solve puzzles presented by such sentences like

(7) Lois believes that Superman flies but does not believe that Clark Kent flies,

and that the "Fregean view", though improved, will have problems with sentences like

Aristotle's mother doubtless believed that Aristotle was Greek,

where it is not clear what modes of presentation of Aristotle and being Greek are actually involved.

To overcome these difficulties, Schiffer presented a theory of belief ascription which he named the *hidden-indexical theory.* The theory ascribes belief as a three-place relation, holding among the believer, the proposition believed, and the mode of presentation under which the believer believes the proposition. The sentence (6) would be represented as

$$\exists m[\Phi^* m \& B(Ralph, \langle Fido, doghood\rangle, m)],$$

where $\Phi^* m$ is "an implicitly referred to and contextually determined type of mode of presentation". Thus by not requiring reference to *particular* modes of presentation, it avoids the pitfall of the Fregean theory. Sentence (7) can be true now because the different types of modes of presentation can be involved in it, one corresponding to 'believes', the other to 'does not believe'.

There are three problems raised by this theory: the candidate problem—whether there are always candidates for modes of presentations, the meaning-intention problem—whether the belief ascribers really mean what the theory requires them to mean, and the logical-form problem—it seems that there is no plausible way to make a sentence ascribing belief into a form that explicitly shows believing is a three-place relation.

To solve the puzzles in a formal way, [1] presented solutions based on the Discourse Representation theory developed by Kamp and Heim, observing that what gives rise to the puzzles is the possible tensions between the so-called internal and "external" connections: a theory capable of dealing with the puzzles should enable one to speak coherently of two distinct beliefs being about the same object in a "internal" sense so that the truth of describing two beliefs as so related does not depend on there being an external object that these two beliefs are about. First, fragments of English are mapped, through an algorithm, onto a system of representations, called discourse representation structures (DRSs); then a mapping is defined from the DRSs into a first order model to determine truth conditions.

A device called an external anchor is used to made explicit the link between a reference marker and a directly referential expression's semantic value. In the case of Kripke's Pierre and London, for example, both reference markers *'Londres'* and 'London' are linked to the city London externally. In the "internal" sense, therefore, Pierre's two beliefs are coherent. Another important component of the structure is what is called familiarity conditions, in the sense that a believer must have familiarity conditions whenever he has a *de re* belief. In the above example, the familiarity conditions which make Pierre hold the two beliefs about *'Londres'* and 'London' are, respectively, having read a book (say) about *'Londres'* and having inhabited in 'London'.

1.2 Philosophical Basis of the Formalization

We develop a logic that formalizes beliefs involving names, and show how to apply it to various philosophical issues relating to belief and referring, including some of the well-known puzzles mentioned above. As far as belief context is concerned, we adopt a basically Frege-Russellian view (as termed in [25]). We take it that in belief contexts, names have senses, the senses are relative to the believers (not always the utterers of the sentences in which the names occur), and modes of presentation constitute senses and are roughly 'identifying descriptions', though the descriptions may pick out nothing, one thing, or more than one thing.

The key of the formalization is the employment—or rather, borrowing—of Frege's notion of modes of presentation. The treatment of modes of presentation here differs from Frege's original notion in certain ways, as will be addressed below.

1.2.1 *De Re* and *De Dicto* Distinction, and Modes of Presentation

We first note that, whereas Kripke [25] dismissed the relevance of *de re* reading of belief attribution to the puzzle he proposed, and thought it's only the *de dicto* reading that was concerned in his article, we do take into account both *de re* and *de dicto* beliefs (so such a formalization of his puzzle will not likely be satisfactory to him). The relevance of both readings will be addressed below.

Our purpose being the formalization of beliefs involving names, modes of presentation (which constitute senses of names) occur only in belief contexts in the formalization. This is not to say that names do not have senses beyond belief contexts, but rather, when *formalizing* sentences, modes of presentation always occur with belief. Like Frege, we regard that in the sentence

'Phosphorus is a planet',

the name 'Phosphorus' has a sense, but the representation of this sentence in our logic system does not involve mode of presentation in its semantics, since no belief is involved. On the other hand, in the sentence

'Paul believes that Phosphorus is a planet',

if we read it as *de dicto*, there is a sense associated with 'Phosphorus' by Paul, and what the sentence says is that Paul believes that Phosphorus, *as he understands it*, is a planet. In the sentence

'Peter believes that Phosphorus is a planet',

'Phosphorus' may have a different sense, which is associated to it by Peter, and the sentence expresses the fact that Peter believes that Phosphorus, as *he* understands it, is a planet.

Thus in the formalization, the sense of a name plays a part in (and only in) (*de dicto*) belief contexts.

In [25] Kripke holds that, if we (who know that Cicero is Tully) say:

'Peter does not believe that Cicero is Tully',

then the names 'Cicero' and 'Tully' should have the senses as *we* understand them, hence, since these two names pick out the same object for us, it is as if saying Peter does not believe that a certain object is not identical to itself, which is absurd, assuming that Peter is a rational person. We, on the other hand, in *de dicto* reading, treat the two names here as having the senses *Peter* understands them.

1.2.2 Relation between Mode of Presentation and the Object

Recall that when Frege introduced the notion of modes of presentation, he talked of the mode of presentation of *that which is designated.* From the term itself, presumably he meant by mode of presentation of 'that which is designated' a certain way in which the object designated presents itself, *e.g.*, some properties of the object. Frege maintained that when using a name in our speech we presuppose a reference, but he did not explain the nature of the mode of presentation associated with the sense of an empty name. Now, if a name does not have a reference (*e.g.*, 'the Golden Mountain'), the mode of presentation associated with its sense is certainly not any properties of 'that which is designated'; but even if a name does have a reference, the mode of presentation associated with its sense still does not need to be any properties of 'that which is designated', since, as it often happens, people may be mistaken in using names to refer things. It is therefore not desirable that there should necessarily be any connection between a mode of presentation and the object designated by the name. Consequently, instead of talking of modes of presentation of *that which is designated*, we talk of modes of presentation *associated with a name.*

It is not unusual that a person does not have any idea about a certain name. For example, if Jones has never heard of the name 'Cicero', then he cannot form any *de dicto* belief about it (thought he may have *de re* belief about it). In that case, Jones does not associate any mode of presentation of 'Cicero'—or the mode of presentation of 'Cicero' for Jones is undefined. As such cases are of little interest for the study of belief and belief reports, we exclude them in our logic; thus, in our logic, given a name, a believer, and a possible world, the mode of presentation of that name for that believer at the given world is always (uniquely) defined.

To summarize our position: belief concerning a name is determined by how the believer understands the name; the way he understands the name is represented by the mode of presentation he associates with the name. Given a name and a believer, there is one and only one mode of presentation associated with the name by the believer at a given moment.

Is there any relation between *de re* and *de dicto* beliefs in which the same name occurs? When can we infer one from the other? We hold that to infer a *de re* belief from a *de dicto* belief concerning the same name, there must need be some connection between the mode of presentation associated with the name for the believer and the object designated by the name. More precisely, if a person 'i' says: 'a is P,' it cannot be inferred that *of* the thing named 'a', i believes that it is P, *unless* the thing named 'a' is indeed what 'i' has in mind when using the name 'a'.

Chapter 2

Formalization: The Logic System FMP

2.1 Overview

The semantics of **FMP** is based on the possible world semantics. A possible world is a state of affairs as conceived possible by some believing agent (as opposed to possible worlds in a metaphysical sense). For a believing agent i, a world Δ is $i-$accessible from a world Γ if, at Γ, i can conceive Δ. At each world there is a domain of objects existent at that world. For technical reason, the domains are subject to the monotonicity condition: if for any believing agent i, Δ is $i-$accessible from Γ, then everything that exists at Γ also exists at Δ. A designation function σ assigns to each name at each world an object: it is possible that the object is not in the domain at that world, which corresponds to the case of an empty name. For each believing agent, each name has a certain mode of presentation at a given world. To understand the motive behind the present way of formalizing modes of presentation, it would be helpful to know the formalization we had attempted in the beginning.

2.1.1 The Motive of the Current Way of Formalizing Modes of Presentation

It seems quite natural to think of modes of presentation as descriptions, hence our initial idea was to formalize modes of presentation as descriptions—that is, as formulas. As a precursor of the present formalization, we had attempted to develop a semantics in which the mode of presentation of a name 'a' for a believing agent i was represented by a formula $m(a, i)$, which has an empty place. For example, if the mode of presentation of the name 'London' for Pierre is the description 'the capital of England', then $m(London, Pierre)$ would be the formula (or predicate, in this case) $C(\, . \,)$, where C is the predicate 'Is Capital of England', the dot in the parenthesis indicating the empty place. We then can have a semantics to the effect that i believes a (as he understands the name 'a') is φ at a world Γ just in case that at Γ, for every object o which satisfies φ, i believes that o is φ. In the example of Pierre, Pierre believes that London is pretty, just in case that he believes that whatever is the capital of England is pretty. The description which a person associates with a name may or may not be uniquely satisfied, but it must be *believed* by that person to uniquely determine an object. For example, a person may think of the name 'Bertrand Russell' as 'the person who wrote *Principia Mathematica*', while unknown to him, the book was co-authored by two people. Or, one may think of 'Unicorn' as a certain animal he saw in the zoo, while in fact the name has no reference. In case someone associate with 'London' the description 'is the capital of France', the description *is* uniquely satisfied, but by Paris instead of London.

2.1.2 The Current Way of Formalizing Modes of Presentation

It was purely for technical reason that we decided to (at least as a first step) represent a mode of presentation by a set of objects, which serves as an abstraction of descriptions: each set can be thought of as consisting of all the objects which satisfy a certain description. A mode of presentation of a name a for a believer i at a possible world Γ is written $m(a, i, \Gamma)$. It may be of any size: in case it contains more than one member, it corresponds to the case when more than one object satisfies a certain description (*e.g.*, "is child of Peter"); in case it is empty ($\emptyset$), it corresponds to cases when no object satisfies the description—this happens when, for example, the description is a contradiction, like "is round and is square"—in such cases, i has a contradictory understanding of the name 'a'.

To say that i has a *de re* belief about a name a (at a world Γ) is to say that the belief i has is about the object designated by a, *viz.*, $\sigma(a, \Gamma)$. On the other hand, *de dicto* beliefs are associated with modes of presentation: i believes that a (as he understands the name) is φ iff he believes that everything that fits the mode of presentation of a for him is φ. Such a semantics thus reflects the point we made in Section 1.2.2, that a mode of presentation associated with a name need not be in any way related to the actual reference of that name, and a connection exists between the two only if the reference of the name fits that mode of presentation; that is, when the reference is a member of the set representing the mode of presentation ($\sigma(a, \Gamma) \in m(a, i, \Gamma)$), for in that case, if i believes that a is φ, then he believes that every member of $m(a, i, \Gamma)$ is φ, thus he believes that $\sigma(a, \Gamma)$ is φ.

We name the logic **FMP** because Frege's ('F') notion of modes of presentation ('MP') plays a central role in its semantics.

2.2 About the Belief Notations

Let i be a believer, a be a name, and φ be a proposition. To distinguish between *de re* and *de dicto* belief, we use two types of expressions for the sentence "i believes that a is φ". The *de re* belief, "Of the thing named 'a', i believes that it is φ", is represented by $B_i\varphi(a)$, whereas the *de dicto* belief, "i believes that a, as he understands it, is φ", is represented by $B_i\langle x.\varphi\rangle(a)$, where x is a variable that serves as a place holder for a (this notation has been inspired by the λ abstract notation which was used in [10] for *de re* scope). The occurrence of 'a' as in $B_i\langle x.\varphi\rangle(a)$ determines that the a has the sense i associates with it, which, in turn, is associated with whatever mode of presentation i associates with a. For a different believer j, the a in $B_i\langle x.\varphi\rangle(a)$ is associated with whatever mode of presentation j associates with a.

For iterated beliefs, as in the sentence "Peter believes that Paul believes that Phosphorus is a planet", we have representations according to whether the name 'Phosphorus' has the sense Peter associates with it or the one Paul associates with it: in the former case, it is represented by $B_{Peter}\langle x.B_{Paul}P(x)\rangle(Ph)$, in the latter case, it is $B_{Peter}B_{Paul}\langle x.P(x)\rangle(Ph)$.

The expressions extend to cases where more than one name is involved, *e.g.*, $B_i\langle xyz.\varphi\rangle(a,b,c)$, $B_i\langle x.B_j\langle y.\varphi\rangle(b)\rangle(a)$, etc.

When the occurrence of a name is not in the scope of the abstract, no sense is associated with it in the formalization: it occurs as *de re*, as the 'a' in $B_i\varphi(a)$.

2.3 The Syntax of FMP

The language $\mathcal{L}_{\mathbf{FMP}}$ consists of the following:

1. a countable set of *constant symbols*: $C = \{a_1, a_2, \ldots\}$;
2. a countable set of *variables*: $V = \{x_1, x_2, \ldots\}$;
3. a countable set $\mathcal{P}$ of predicates: let $\mathcal{P}^n$ denote the set of $n-$place predicates, where $n = 1, 2, \ldots$, with $E \in \mathcal{P}^1$ as the existential predicate. Then $\mathcal{P} = \bigcup_n \mathcal{P}^n$;
4. the equality sign "=".

Remark The constant symbols are what we refer to as *names* in less formal contexts.‘

A *term* is a constant or a variable (note that we do not have function symbols).

Definition 2.3.1 *Atomic formulas* are all expressions of the form $P(t_1, \ldots, t_n)$ or $t_1 = t_2$, where $P \in \mathcal{P}$, $\sqcup_\infty, \ldots, \sqcup_\backslash$ are terms.

The set Φ of *formulas* and the set Ψ of *pseudo-formulas* are defined by a mutual recursion:

1. Φ is the smallest set such that

   (1) all atomic formulas are in Φ; and

(2) if $\varphi_1, \varphi_2 \in \Phi$, and $\psi \in \Psi$, then

$\neg\varphi_1$,

$\varphi_1 \wedge \varphi_2$,

$(\exists x)\varphi_1$,

$B_i\psi$

are in Φ, where $i \in I = \{1, 2, \ldots\}$, n is a positive integer, x is a variable.

2. Ψ is the smallest set such that

   (1) $\Phi \subseteq \Psi$; and

   (2) if $\psi \in \Psi$, $x \in V$, $a \in C$, x has not "represented" any constant in ψ, and a has not "acquired a sense" in ψ, then $\langle x.\psi\rangle(a) \in \Psi$—where by "representing a constant" and "acquiring a sense" we mean the following:

   i. for any formula $\varphi \in \Phi$, no variable represents any constant in φ, and no constant acquires a sense in φ;

   ii. for $\psi \in \Psi$, the variables which have represented constants in $\langle x.\psi\rangle(a)$ are those which have represented constants in ψ, together with x, and the constants which have acquired senses in $\langle x.\psi\rangle(a)$ are those which have acquired senses in ψ, together with a.

Definition 2.3.2 $\langle x_1 \ldots x_n.\varphi\rangle(a_1, \ldots, a_n)$ is an abbreviation for the pseudo-formula $\langle x_1.\langle x_2.\ldots\langle x_n.\varphi\rangle(a_n)\rangle\ldots(a_2)\rangle(a_1)$. We call the occurrences of the $a_1, \ldots, a_n$ in such positions *Fregean occurrences*.

Remark 1. $\vec{x}$ abbreviates $x_1, \ldots, x_n$, and $\vec{a}$ abbreviates $a_1, \ldots, a_n$.

2. For readability, $\varphi(\vec{t})$ abbreviates $\varphi(t_1/x_1, \ldots, t_n/x_n)$, which is the result of replacing all free occurrences of x_j in φ by occurrences of $t_j (j = 1, \ldots, n)$.

3. In certain contexts we write, *e.g.*, $B_i\langle x.\varphi(x,t)\rangle(a)$ instead of $B_i\langle x.\varphi\rangle(a)$ so as to make distinction between $B_i\langle x.\varphi(x,t)\rangle(a)$ and, say, $B_i\langle x.\varphi(x,y)\rangle(a)$.

Definition 2.3.3 [nesting complexity] The *nesting complexity* of a pseudo-formula ψ is defined recursively as follows:

1. If $\psi \in \Phi$, then the nesting complexity of ψ is 0;
2. If ψ is $\langle x.\varphi\rangle(a)$, where φ is of nesting complexity n, then the nesting complexity of ψ is $n + 1$.

Definition 2.3.4 [$\theta-$formula] A *$\theta-$formula* is any formula of the form $\neg B_i\langle x.\psi\rangle(a)$ for some constant a and some formula ψ.

Remark
1. The name '$\theta-$formula' was an imitation of the '$\delta-$ formula' used by Smullyan (*e.g.*, in [58]) and Fitting (*e.g.*, in [8]).
2. The symbols θ or $\theta_1, \theta_2, \ldots$ will also be used to denote individual θ-formulas.

Definition 2.3.5 Let $\theta = \neg B_i\langle x.\psi\rangle(a)$, $\mathcal{L}$ be a sublanguage of $\mathcal{L}_{\mathbf{FMP}}$, and z be a variable, then

$$\theta^{\mathcal{L}}(z) \overset{df}{=} \bigcup_{\varphi'}\{(\forall\vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(z)) \land E(z) \land \neg B_i\psi(z)\},$$

where φ' is a formula of $\mathcal{L}$ (thus $\bigcup_{\varphi'}$ is a union over all formulas of $\mathcal{L}$), and $\vec{y}$ includes all the variables of $\mathcal{L}$ which occur in φ'.

2.4 The Semantics of FMP

Definition 2.4.1 [**FMP** model] A **FMP** *model* is a tuple

$$M = \langle G, I, R_1, \ldots, R_{|I|}, D, \sigma, m, \pi\rangle,$$

where

1. G is a set of possible worlds.
2. I is the set of believers.

3. $R_i \subseteq G^2$, for each $i \in I$, is the accessibility relation which is serial, transitive and Euclidean (this condition can be relaxed, and the system can be modified to deal with other propositional attitudes such as knowledge by replacing this conditions accordingly).

4. For each $\Gamma \in G$, $D(\Gamma)$ is the domain at Γ. D is monotonic: if $\Gamma R_i \Delta$, then $D(\Gamma) \subseteq D(\Delta)$.

   For convenience, we shall refer $\bigcup_{\Gamma \in G} D(\Gamma)$, the union of the domains of all worlds of M, as the *universe* of M.

5. $\sigma : C \times G \mapsto \bigcup_{\Gamma \in G} D(\Gamma)$, the *designation function*, assigns to each name at each possible world an object: $\sigma(a, \Gamma) \in \bigcup_{\Gamma \in G} D(\Gamma)$. σ is subject to the *local rigidity* condition: for any constant a, any Γ, and any Δ such that $\Gamma R_i \Delta$ for some i, $\sigma(a, \Gamma) = \sigma(a, \Delta)$.

6. $m : C \times I \times G \mapsto 2^{\bigcup_{\Gamma \in G} D(\Gamma)}$, the *sense function*, is a function such that for each $a \in C, i \in I, \Gamma \in G$, $m(a, i, \Gamma) \subseteq D(\Gamma)$. This function represents the *mode of presentation* of the name a for the believer i at the world Γ. It's possible that $m(a, i, \Gamma) = \emptyset$.

7. For each $\Gamma \in G$ and each $n = 1, 2, \ldots$, $\pi : (\mathcal{P}^n - \{E\}) \times \Gamma \mapsto 2^{\bigcup_{\Omega \in G} D(\Omega)^n}$ is the interpretation of atomic formulas at each world.

Definition 2.4.2 Let $M = \langle G, I, R_1, \ldots, R_{|I|}, D, \sigma, m, \pi \rangle$ be a **FMP** model. An *assignment* μ is a mapping: $V \mapsto \bigcup_{\Omega \in G} D(\Omega)$.

Using the same symbol μ, we define, for each $x \in V$, $a \in C$, $\Gamma \in G$:

$$\mu(x, \Gamma) \overset{df}{=} \mu(x),$$

$$\mu(a, \Gamma) \overset{df}{=} \sigma(a, \Gamma).$$

Definition 2.4.3 [satisfaction] The *satisfaction* relation $\models$ is defined as follows.

Atom For atomic formulas: if $t_1, \ldots, t_n$ are terms, then

$$\Gamma, \mu \models P(t_1, \ldots, t_n) \text{ iff } \langle \mu(t_1, \Gamma), \ldots, \mu(t_n, \Gamma) \rangle \in \pi(P, \Gamma), \text{ where } P \in \mathcal{P}^n - \{E\};$$

$\Gamma, \mu \models E(t)$ iff $\mu(t, \Gamma) \in D(\Gamma)$;

$\Gamma, \mu \models t_1 = t_2$ iff $\mu(t_1, \Gamma) = \mu(t_2, \Gamma)$.

Neg $\Gamma, \mu \models \neg\varphi$ iff $\Gamma, \mu \not\models \varphi$.

Conj $\Gamma, \mu \models \varphi \wedge \psi$ iff $\Gamma, \mu \models \varphi$ and $\Gamma, \mu \models \psi$.

Quant $\Gamma, \mu \models (\exists x)\varphi$ iff $\Gamma, \mu \models \varphi[x : o]$ for some $o \in D(\Gamma)$.

B-r $\Gamma, \mu \models B_i\varphi$ iff for every Δ such that $\Gamma R_i \Delta$, $\Delta, \mu \models \varphi$,

where $\varphi \in \Phi$.

B-d $\Gamma, \mu \models B_i\langle x.\psi\rangle(a)$ iff for every $o \in m(a, i, \Gamma)$, $\Gamma, \mu \models B_i\psi[x : o]$,

Remark 1. The notation "$\Gamma, \mu \models \varphi[x : o]$" is a shorthand for:

"$\Gamma, \mu' \models \varphi$, where μ' is an assignment like μ except that it assigns the object o to the free variable x".

2. Because of the condition met by σ, all constants are (locally) rigid; as a result, **B-r** gives $B_i\phi(a)$ (where $\phi(a) \in \Phi$) a *de re* interpretation, as opposed to the *de dicto* interpretation of $B_i\langle x.\psi\rangle(a)$.

Proposition 2.4.4 $\Gamma, \mu \models B_i\langle x_1 \ldots x_n.\varphi\rangle(a_1, \ldots, a_n)$
iff $\Gamma, \mu \models B_i\langle x_{k_1} \ldots x_{k_n}.\varphi\rangle(a_{k_1}, \ldots, a_{k_n})$, *where* $\langle k_1, \ldots, k_n\rangle$ *is any permutation of* $\langle 1, \ldots, n\rangle$
iff for every $o_1 \in m(a_1, i, \Gamma), \ldots, o_n \in m(a_n, i, \Gamma)$, $\Gamma, \mu \models B_i\varphi[x_1 : o_1; \ldots ; x_n : o_n]$
*(*cf. *Definition 2.3.2).*

Definition 2.4.5 A formula φ is *valid* in a **FMP** model M, written $M \models \varphi$, iff for every assignment μ, $M, \mu \models \varphi$. φ is **FMP** *valid*, written $\models \varphi$, iff it is valid in every **FMP** model.

2.5 Axioms and Rules of Inference

In the following, $\varphi, \psi \in \Phi$, a, b are constants, and t_1, t_2 are terms, unless indicated otherwise.

2.5.1 Axioms

FMP 0 All tautologies

FMP 1 Substitution Principle: $t_1 = t_2 \supset (\varphi \supset \psi)$, where ψ is a formula resulting from free substitution [22] of one or more occurrences of t_1 in the formula φ by t_2, with the restriction that, if t_1 is a constant, then all the occurrences of t_1 in φ which are being substituted by t_2 are non-Fregean occurrences (cf. Definition 2.3.2).

FMP 2 $(\forall x)E(x)$

FMP 3 $(\forall x)\varphi(x) \wedge E(t) \supset \varphi(t)$

FMP 4 $x = x$

FMP b1 $(\forall x)B_iE(x)$

FMP b2 $B_i(\varphi \supset \psi) \supset (B_i\varphi \supset B_i\psi)$

FMP b3 $B_i\varphi \supset \neg B_i\neg\varphi$

FMP b4 $B_i\varphi \supset B_iB_i\varphi$

FMP b5 $\neg B_i\varphi \supset B_i\neg B_i\varphi$

FMP b6a $x = y \supset B_i(x = y)$

FMP b6b $x \neq y \supset B_i(x \neq y)$

FMP bs1 $B_i\langle xy.\varphi\rangle(a, b) \equiv (\forall x)\{(\forall y)[B_i\langle x.\varphi\rangle(a) \supset B_i\varphi] \supset B_i\langle y.\varphi\rangle(b)\}$, where $\varphi \in \Psi$

FMP bs2 $\theta \supset (\exists w) \bigwedge \theta_0^{\mathcal{L}}(w)$,

for *any* finite subset $\theta_0^{\mathcal{L}}(w)$ of $\theta^{\mathcal{L}}(w)$, where $\bigwedge \theta_0^{\mathcal{L}}(w)$ is the conjunction of all members of $\theta_0^{\mathcal{L}}(w)$, w is a variable not in the sublanguage $\mathcal{L}$of $\mathcal{L}_{\mathbf{FMP}}$

Remark 1. Intuitively, **FMP bs1** translates as, for example, Smith believes that Cicero killed Caesar (as he understands these two names), just in case that, for anyone x

such that whenever Smith believes that Cicero killed any person, then he believes that x killed the same person for any such x, Smith believes that x killed Caesar.

2. For **FMP bs2**, recall the definition of $\theta^{\mathcal{L}}(w)$ as given in Definition 2.3.5: if $\theta = \neg B_i\langle x.\psi\rangle(a)$, then

$$\theta^{\mathcal{L}}(w) = \bigcup_{\varphi'}\{(\forall\vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w)) \wedge E(w) \wedge \neg B_i\psi(w)\},$$

where $\vec{y}$ includes all the variables of $\mathcal{L}$ and φ' is a formula of $\mathcal{L}$.

3. Since the axioms include **FMP b3**, **FMP b4** and **FMP b5**, this is a **K45** system.

2.5.2 Rules of Inference

R1 Modus Ponens:

$$\frac{\varphi,\ \varphi \supset \psi}{\psi}$$

R2 Universal Generalization:

$$\frac{\varphi \supset \psi(z)}{\varphi \supset (\forall w)\psi(w)}$$

where $\psi(z)$ is the result of replacing all free occurrences of w in $\psi(w)$ by occurrences of z, and z does not occur free in φ or $\psi(w)$.

R3 *De Re* Belief Generalization:

$$\frac{\varphi}{B_i\varphi}$$

R4

$$\frac{\varphi(\vec{x})}{\varphi(\vec{a})}$$

where $\vec{x}$ occur free in φ, and $\varphi(\vec{a})$ is the result of replacing $\vec{x}$ by $\vec{a}$.

2.6 Derived Rules of Inference and Some Theorems of FMP

Derived Rule DR5 *De Re* Belief Distribution:

$$\frac{B_i(\varphi \supset \psi)}{B_i\varphi \supset B_i\psi}$$

Proof Derivable from Axiom **FMP b2** and **R1**. ■

Theorem 2.6.1 $\vdash B_i(\varphi \wedge \psi) \equiv B_i\varphi \wedge B_i\psi$.

Proof (1) (Left to right) Follows from propositional tautologies, **R1** and Axiom **FMP b2**.

(2) (Right to left)

$\vdash \varphi \supset (\psi \supset (\varphi \wedge \psi))$ (tautology),

$\vdash B_i[\varphi \supset (\psi \supset (\varphi \wedge \psi))]$ (**R3**),

$\vdash B_i\varphi \supset B_i[\psi \supset (\varphi \wedge \psi)]$ (Derived rule **DR5**),

$\vdash B_i\varphi \supset [B_i\psi \supset B_i(\varphi \wedge \psi)]$ (Axiom **FMP b2**),

$\vdash B_i\varphi \wedge B_i\psi \supset B_i(\varphi \wedge \psi)$. ■

Theorem 2.6.2 *(1)* $\vdash a = b \supset B_i(a = b)$;

(2) $\vdash a \neq b \supset B_i(a \neq b)$.

Proof Follows directly from Axioms **FMP b6a**, **FMP b6b**, and **R4**. ■

Theorem 2.6.3 $\vdash (\forall x)B_i\varphi \supset B_i\langle x.\varphi\rangle(a)$.

Proof $\vdash \neg B_i\langle x.\varphi\rangle(a) \supset (\exists w)\{\bigwedge_{\varphi' \in S}[(\forall \vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w))] \wedge E(w) \wedge \neg B_i\varphi(w)\}$,

for any finite set S of formulas φ' of $\mathcal{L}_{\mathbf{FMP}}$ (Axiom **FMP bs2** and Definition 2.3.5),

$\vdash \neg B_i\langle x.\varphi\rangle(a) \supset (\exists w)\neg B_i\varphi(w)$ (first-order tautologies);

taking its contrapositive, and replacing the bounded variable w by x, we have

$\vdash (\forall x)B_i\varphi(x) \supset B_i\langle x.\varphi\rangle(a)$. ■

Corollary 2.6.4 $\vdash (\forall \vec{x})B_i\varphi \supset B_i\langle \vec{x}.\varphi\rangle(\vec{a})$.

Proof Follows from Axiom **FMP bs1** and Theorem 2.6.3. ■

Derived Rule DR6

$$\frac{B_i\varphi}{B_i\langle x.\varphi\rangle(a)}$$

where $\varphi \in \Phi$.

Proof Derivable from **R2** and Theorem 2.6.3. ■

Derived Rule DR7 *De Dicto* Belief Generalization:

$$\frac{\varphi}{B_i\langle x.\varphi\rangle(a)}$$

Proof Derivable from **R3** and Derived Rule **DR6**. ■

Derived Rule DR8

$$\frac{\bigwedge \theta_0^{\mathcal{L}}(z) \supset \chi}{\theta \supset \chi}$$

where z is a variable not in $\mathcal{L}$, $\theta_0^{\mathcal{L}}(z)$ is any finite subset of $\theta^{\mathcal{L}}(z)$, and $\bigwedge \theta_0^{\mathcal{L}}(z)$ is the conjunction of all members of $\theta_0^{\mathcal{L}}(z)$.

Proof Derivable from **R2** and **FMP bs2**:

Let $\theta_0^{\mathcal{L}}(z)$ be any finite subset of $\theta^{\mathcal{L}}(z)$, and $\bigwedge \theta_0^{\mathcal{L}}(z)$ be the conjunction of all the members of $\theta_0^{\mathcal{L}}(z)$.

Since z is not in $\mathcal{L}$, it does not occur (hence not occur free) in χ (therefore $\neg\chi$) or any formula φ' of $\mathcal{L}$ contained in $\theta_0^{\mathcal{L}}(w)$, so by **R2**,

$$\vdash \neg\chi \supset \neg \bigwedge \theta_0^{\mathcal{L}}(z) \Longrightarrow \vdash \neg\chi \supset (\forall w)\neg \bigwedge \theta_0^{\mathcal{L}}(w) \tag{2.1}$$

By **FMP bs2**, taking its contrapositive form,

$$\vdash (\forall w)(\neg \bigwedge \theta_0^{\mathcal{L}}(w)) \supset \neg\theta \tag{2.2}$$

Combining (2.1) and (2.2),

$\vdash \neg\chi \supset \neg \bigwedge \theta_0^{\mathcal{L}}(z) \Longrightarrow \vdash \neg\chi \supset \neg\theta$, that is,

$\vdash \bigwedge \theta_0^{\mathcal{L}}(z) \supset \chi \Longrightarrow \vdash \theta \supset \chi$.

■

Theorem 2.6.5 $\vdash B_i\langle x.\varphi \supset \psi\rangle(a) \supset (B_i\langle x.\varphi\rangle(a) \supset B_i\langle x.\psi\rangle(a))$.

Proof By Axiom **FMP bs2** and Definition 2.3.5,

$\vdash B_i\langle x.\varphi\rangle(a) \wedge \neg B_i\langle x.\psi\rangle(a) \supset$

$B_i\langle x.\varphi\rangle(a) \wedge (\exists w)\{\bigwedge_{\varphi' \in S}[(\forall \overline{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w))] \wedge E(w) \wedge \neg B_i\psi(w)\}$,

for any finite $S \subset \Phi$.

Since $\varphi \in \Phi$, it is in some such finite set S, so it follows that

$\vdash B_i\langle x.\varphi\rangle(a) \wedge \neg B_i\langle x.\psi\rangle(a) \supset$

$(\exists w)\{\bigwedge_{\varphi' \in S}[(\forall \overline{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w))] \wedge E(w) \wedge B_i\varphi(w) \wedge \neg B_i\psi(w)\}$,

for any finite $S \subset \Phi$, *i.e.*,

$\vdash B_i\langle x.\varphi\rangle(a) \wedge \neg B_i\langle x.\psi\rangle(a) \supset$

$(\exists w)\{\bigwedge_{\varphi' \in S}[(\forall \overline{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w))] \wedge E(w) \wedge \neg(B_i\varphi(w) \supset B_i\psi(w))\}$,

for any finite $S \subset \Phi$; by Axiom **FMP b2**,

$\vdash B_i\langle x.\varphi\rangle(a) \wedge \neg B_i\langle x.\psi\rangle(a) \supset$

$(\exists w)\{\bigwedge_{\varphi' \in S}[(\forall \overline{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w))] \wedge E(w) \wedge \neg B_i(\varphi(w) \supset \psi(w))\}$,

for any finite $S \subset \Phi$, *i.e.*,

$\vdash B_i\langle x.\varphi\rangle(a) \wedge \neg B_i\langle x.\psi\rangle(a) \supset$

$(\exists w)\{\bigwedge_{\varphi' \in S}[(\forall \overline{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w))] \wedge E(w) \wedge \neg B_i(\varphi \supset \psi)(w)\}$,

for any finite $S \subset \Phi$.

Now, again, since $\varphi(w) \supset \psi(w) \in \Phi$, it is in some finite $S \subset \Phi$, hence it follows that

$\vdash B_i\langle x.\varphi\rangle(a) \wedge \neg B_i\langle x.\psi\rangle(a) \supset \neg B_i\langle x.\varphi \supset \psi\rangle(a)$, therefore

$\vdash B_i\langle x.\varphi \supset \psi\rangle(a) \supset (B_i\langle x.\varphi\rangle(a) \supset B_i\langle x.\psi\rangle(a))$. ■

Derived Rule DR9 *De Dicto* Belief Distribution:

$$\frac{B_i\langle x.\varphi \supset \psi\rangle(a)}{B_i\langle x.\varphi\rangle(a) \supset B_i\langle x.\psi\rangle(a)}$$

Proof Derivable from Theorem 2.6.5 and **R1**. ■

Theorem 2.6.6 $B_i\langle x.\varphi \wedge \psi\rangle(a) \equiv B_i\langle x.\varphi\rangle(a) \wedge B_i\langle x.\psi\rangle(a)$

Proof (1) (Left to right) Follows directly from Theorem 2.6.5.

(2) (Right to left) Similar reasoning as in the proof of Theorem 2.6.1:

$\vdash \varphi \supset (\psi \supset (\varphi \wedge \psi))$ (tautology),

$\vdash B_i\langle x.\varphi \supset (\psi \supset (\varphi \wedge \psi))\rangle(a)$ (Derived Rule **DR7**),

$\vdash B_i\langle x.\varphi\rangle(a) \supset B_i\langle x.\psi \supset (\varphi \wedge \psi)\rangle(a)$ (Derived Rule **DR9**),

$\vdash B_i\langle x.\varphi\rangle(a) \supset [B_i\langle x.\psi\rangle(a) \supset B_i\langle x.\varphi \wedge \psi\rangle(a)]$ (Theorem 2.6.5),

$\vdash B_i\langle x.\varphi\rangle(a) \wedge B_i\langle x.\psi\rangle(a) \supset B_i\langle x.\varphi \wedge \psi\rangle(a)$. ■

Theorem 2.6.7 $\vdash B_i\langle x.\varphi\rangle(a) \supset B_i\langle x.B_i\varphi\rangle(a)$, *where* $\varphi \in \Psi$

Proof $\vdash \neg B_i\langle x.B_i\varphi\rangle(a) \supset$

$(\exists w)\{\bigwedge_{\varphi' \in S}[(\forall \vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w))] \wedge E(w) \wedge \neg B_i B_i\varphi(w)\}$, * for any finite $S \subset \Phi$ (Axiom **FMP bs2** and Definition 2.3.5),

$\vdash \neg B_i\langle x.B_i\varphi\rangle(a) \supset$

$(\exists w)\{\bigwedge_{\varphi'}[(\forall \vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w))] \wedge E(w) \wedge \neg B_i\varphi(w)\}$,

for any finite $S \subset \Phi$ (contrapositive of Axiom **b4**);

since $\varphi \in \Phi$, it is in some finite $S \subset \Phi$, hence it follows that

$\vdash \neg B_i\langle x.B_i\varphi\rangle(a) \supset \neg B_i\langle x.\varphi\rangle(a)$. ■

Lemma 2.6.8 (Converse Barcan Formula) $\vdash B_i(\forall x)\varphi \supset (\forall x)B_i\varphi$.

Proof $\vdash (\forall x)\varphi(x) \wedge E(t) \supset \varphi(t)$ (**FMP 3**),

$\vdash (\forall x)\varphi(x) \supset (E(t) \supset \varphi(t))$,

$\vdash B_i[(\forall x)\varphi(x) \supset (E(t) \supset \varphi(t))]$ (**R3**),

$\vdash B_i(\forall x)\varphi(x) \supset B_i(E(t) \supset \varphi(t))$ (Derived Rule **DR5**),

$\vdash B_i(\forall x)\varphi(x) \supset (\forall x)B_i(E(x) \supset \varphi(x))$ (**R2**),

$\vdash B_i(\forall x)\varphi(x) \supset (\forall x)(B_iE(x) \supset B_i\varphi(x))$ (**FMPb2**),

$\vdash B_i(\forall x)\varphi(x) \supset [(\forall x)B_iE(x) \supset (\forall x)B_i\varphi(x)]$ (tautology),

$\vdash B_i(\forall x)\varphi(x) \land (\forall x)B_iE(x) \supset (\forall x)B_i\varphi(x)$,

using **FMP b1**, $\vdash (\forall x)B_iE(x)$, it follows that

$\vdash B_i(\forall x)\varphi(x) \supset (\forall x)B_i\varphi(x)$. ■

2.7 Soundness

Theorem 2.7.1 (Soundness) *The axiom system of* **FMP** *is sound with respect to the* **FMP** *semantics.*

Proof (1) Each of the axioms is valid: (the proofs in most classical cases are omitted)

FMP b1 Follows directly from the monotonicity of D.

FMP b3 Assume the contrary: there is some Γ of some model M and some μ such that

$\Gamma, \mu \models B_i\varphi \land B_i\neg\varphi$, then for every Δ such that $\Gamma R_i \Delta$,

$\Delta, \mu \models \varphi$ and $\Delta, \mu \models \neg\varphi$ (by **B-r**), so

$\Delta, \mu \models \varphi \land \neg\varphi$,

which is a contradiction.

FMP b4 Suppose $\Gamma, \mu \models B_i\varphi$. Then by **B-r**, for all Δ such that $\Gamma R_i \Delta$, $\Delta, \mu \models \varphi$.

By the transitivity property of R_i, for every Σ such that $\Delta R_i \Sigma$, $\Gamma R_i \Sigma$, so the above holds for any such Σ as well: $\Sigma, \mu \models \varphi$;

so by **B-r** again, $\Delta, \mu \models B_i\varphi$;

thus, by **B-r** again, $\Gamma, \mu \models B_iB_i\varphi$.

Hence $\Gamma, \mu \models B_i\varphi \supset B_iB_i\varphi$.

FMP b5 Suppose $\Gamma \models \neg B_i\varphi$. Then by **B-r**, for some Δ such that $\Gamma R_i \Delta$, $\Delta, \mu \models \neg\varphi$.

By the Euclidean property of R_i, for every Σ such that $\Gamma R_i \Sigma$, $\Sigma R_i \Delta$, therefore by the above and **B-r**, $\Sigma, \mu \models \neg B_i\varphi$;

therefore, by **B-r** again, $\Gamma, \mu \models B_i \neg B_i \varphi$.

Therefore $\neg B_i\varphi \supset B_i\neg B_i\varphi$.

FMP b6a, FMP b6b Follow directly from **Atom** and **B-r** and the fact that variables are rigid and constants are locally rigid designators.

FMP bs1 ($\Longrightarrow$)

Assume the contrary, that is, there is some Γ in some **FMP** model M such that for some assignment μ,

$\Gamma, \mu \models B_i\langle xy.\varphi\rangle(a,b) \wedge \neg(\forall x)\{(\forall y)[B_i\langle x.\varphi\rangle(a) \supset B_i\varphi] \supset B_i\langle y.\varphi\rangle(b)\}$, then

$$\Gamma, \mu \models B_i\langle xy.\varphi\rangle(a,b) \tag{2.3}$$

and

$$\Gamma, \mu \models (\exists x)\{(\forall y)[B_i\langle x.\varphi\rangle(a) \supset B_i\varphi] \wedge \neg B_i\langle y.\varphi\rangle(b)\} \tag{2.4}$$

By (2.3), for every $o_1 \in m(a,i,\Gamma)$, every $o_2 \in m(b,i,\Gamma)$,

$$\Gamma, \mu \models B_i\varphi[x : o_1; y : o_2] \tag{2.5}$$

By (2.4), there is some $\hat{o} \in D(\Gamma)$ such that

$\Gamma, \mu \models (\forall y)[B_i\langle x.\varphi\rangle(a) \supset B_i\varphi] \wedge \neg B_i\langle y.\varphi\rangle(b)[x : \hat{o}]$, that is,

for every $o \in D(\Gamma)$,

$$\Gamma, \mu \models B_i\langle x.\varphi\rangle(a) \supset B_i\varphi[x : \hat{o}; y : o] \tag{2.6}$$

and

$$\Gamma, \mu \models \neg B_i \langle y.\varphi \rangle (b)[x : \hat{o}] \tag{2.7}$$

The latter (2.7) means that (for every $o \in D(\Gamma)$) there is some $\hat{o}_2 \in m(b, i, \Gamma)$ such that

$$\Gamma, \mu \models \neg B_i \varphi [x : \hat{o}; y : \hat{o}_2] \tag{2.8}$$

The former (2.6) means for every $o \in D(\Gamma)$,

$\Gamma, \mu \models \neg B_i \langle x.\varphi \rangle (a) \vee B_i \varphi [x : \hat{o}; y : o]$, that is,

for every $o \in D(\Gamma)$, either

there is some $\hat{o}_1 \in m(a, i, \Gamma)$ such that

$$\Gamma, \mu \models \neg B_i \varphi [x : \hat{o}_1; y : o] \tag{2.9}$$

or

$$\Gamma, \mu \models B_i \varphi [x : \hat{o}; y : o] \tag{2.10}$$

Since $\hat{o}_2 \in m(b, i, \Gamma) \subseteq D(\Gamma)$, it can substitute for o in the above, that is, either (2.9) or (2.10) should hold when o is substituted by $\hat{o}_2$.

[Case 1] If (2.9) holds with o replaced by $\hat{o}_2$, then there is some $\overline{o}_1 \in m(a, i, \Gamma)$ such that

$\Gamma, \mu \models \neg B_i \varphi [x : \overline{o}_1; y : \hat{o}_2]$, which contradicts (2.5).

[Case 2] If (2.10) holds with o replaced by $\hat{o}_2$, then

$\Gamma, \mu \models B_i \varphi [x : \hat{o}; y : \hat{o}_2]$, which contradicts (2.8).

Thus in either case a contradiction results.

$(\Longleftarrow)$

Assume the contrary, that is, there is some Γ in some **FMP** model M such that for some assignment μ,

$\Gamma, \mu \models (\forall x)\{(\forall y)[B_i\langle x.\varphi\rangle(a) \supset B_i\varphi] \supset B_i\langle y.\varphi\rangle(b)\} \wedge \neg B_i\langle xy.\varphi\rangle(a, b)$,

that is,

$$\Gamma, \mu \models (\forall x)\{(\forall y)[B_i\langle x.\varphi\rangle(a) \supset B_i\varphi] \supset B_i\langle y.\varphi\rangle(b)\} \tag{2.11}$$

and

$$\Gamma, \mu \models \neg B_i\langle xy.\varphi\rangle(a, b) \tag{2.12}$$

The latter (2.12) means there is some $\hat{o}_1 \in m(a, i, \Gamma)$, some $\hat{o}_2 \in m(b, i, \Gamma)$ such that

$$\Gamma, \mu \models \neg B_i\varphi[x : \hat{o}_1; y : \hat{o}_2] \tag{2.13}$$

The former (2.11) means that for every $o \in D(\Gamma)$,

$\Gamma, \mu \models (\forall y)[B_i\langle x.\varphi\rangle(a) \supset B_i\varphi] \supset B_i\langle y.\varphi\rangle(b)[x : o]$, that is,

for every $o \in D(\Gamma)$,

$\Gamma, \mu \models \neg(\forall y)[B_i\langle x.\varphi\rangle(a) \supset B_i\varphi] \vee B_i\langle y.\varphi\rangle(b)[x : o]$, that is,

for every $o \in D(\Gamma)$, either

$$\Gamma, \mu \models (\exists y)\neg[B_i\langle x.\varphi\rangle(a) \supset B_i\varphi][x : o] \tag{2.14}$$

or

$$\Gamma, \mu \models B_i\langle y.\varphi\rangle(b)[x : o] \tag{2.15}$$

The latter (2.15) means for every $o_2 \in m(b, i, \Gamma)$,

$$\Gamma, \mu \models B_i\varphi[x : o; y : o_2] \tag{2.16}$$

The former (2.14) means that

$\Gamma, \mu \models (\exists y)[B_i\langle x.\varphi\rangle(a) \wedge \neg B_i\varphi][x : o]$, *i.e.*,

there is some $o' \in D(\Gamma)$ such that

$\Gamma, \mu \models B_i\langle x.\varphi\rangle(a) \wedge \neg B_i\varphi[x : o; y : o']$, *i.e.*,

for every $o_1 \in m(a, i, \Gamma)$,

$$\Gamma, \mu \models B_i\varphi[x : o_1; y : o'] \tag{2.17}$$

and

$$\Gamma, \mu \models \neg B_i\varphi[x : o; y : o'] \tag{2.18}$$

So far we have:

1) (2.13) holds, and

2) for every $o \in D(\Gamma)$, either both (2.17) and (2.18) hold, or (2.16) holds.

Now take the $\hat{o}_1 \in m(a, i, \Gamma) \subseteq D(\Gamma)$ in (2.13). Since it is a member of $D(\Gamma)$, by 2) above, substituting it for o, it should either satisfy both (2.17) and (2.18), or satisfy (2.16). But because of (2.13), $\hat{o}_1$ cannot satisfy (2.16); therefore it must satisfy both (2.17) and (2.18). Hence we should have:

there is some o^1 such that for every $o_1 \in m(a, i, \Gamma)$,

$$\Gamma, \mu \models B_i\varphi[x : o_1; y : o^1] \tag{2.19}$$

(from (2.17)) and

$$\Gamma, \mu \models \neg B_i\varphi[x : \hat{o}_1; y : o^1] \tag{2.20}$$

(from (2.18)).

Letting the o_1 in (2.19) be $\hat{o}_1$, we get

$\Gamma, \mu \models B_i\varphi[x : \hat{o}_1; y : o^1]$, which contradicts (2.20).

Therefore the original assumption is false, and the axiom is valid.

FMP bs2 Assume the contrary, *i.e.*, there is some model M, some possible world Γ in M and some assignment μ such that for some finite subset $\theta_0^{\mathcal{L}}(w)$ of $\theta^{\mathcal{L}}(w)$,

$$M, \Gamma, \mu \models \theta \wedge \neg(\exists w) \bigwedge \theta_0^{\mathcal{L}}(w)$$

(where $\bigwedge \theta_0^{\mathcal{L}}(w)$ is the conjunction of the members of $\theta_0^{\mathcal{L}}(w)$), *i.e.*,

$$M, \Gamma, \mu \models (\forall w)[\bigwedge_{\varphi'} (\forall \vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w)) \wedge E(w) \supset B_i\psi(w)] \tag{2.21}$$

and

$$M, \Gamma, \mu \models \neg B_i\langle x.\psi\rangle(a) \tag{2.22}$$

where $\bigwedge_{\varphi'}$ is the conjunction taken over the *finite* subset of formulas φ' of $\mathcal{L}$ which is part of $\theta_0^{\mathcal{L}}(w)$.

By (2.22), there is some $\hat{o} \in m(a, i, \Gamma) \subseteq D(\Gamma)$ such that

$$M, \Gamma, \mu \models \neg B_i\psi[x : \hat{o}] \tag{2.23}$$

By (2.21), for every $o \in D(\Gamma)$,

$M, \Gamma, \mu \models \bigwedge_{\varphi'} (\forall \vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w)) \wedge E(w) \supset B_i\psi(w)[w : o]$;

let o be substituted by $\hat{o}$, we have

$$M, \Gamma, \mu \models \bigwedge_{\varphi'} (\forall \vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w)) \wedge E(w) \supset B_i\psi(w)[w : \hat{o}] \tag{2.24}$$

Since $\hat{o} \in m(a, i, \Gamma) \subseteq D(\Gamma)$,

$M, \Gamma, \mu \models E(w)[w : \hat{o}]$ and

for every φ' in any finite subset of formulas of $\mathcal{L}$, and every $o_1, o_2, \ldots \in D(\Gamma)$, whenever $M, \Gamma, \mu \models B_i\langle x.\varphi'\rangle(a)[y_1 : o_1; y_2 : o_2; \ldots]$, then $M, \Gamma, \mu \models B_i\varphi'[x : \hat{o}; y_1 : o_1; y_2 : o_2; \ldots]$ (where $y_1, y_2, \ldots$ are the elements of $\vec{y}$), so

$M, \Gamma, \mu \models \bigwedge_{\varphi'} (\forall \vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(w)) \wedge E(w)[w : \hat{o}]$;

combining this with (2.24),

$M, \Gamma, \mu \models B_i\psi(w)[w : \hat{o}]$,

contradictory to (2.23).

(2) Each of the rules of inferences preserves validity.

R2 Suppose $\varphi \supset (\forall w)\psi(w)$ is not valid, then there is some model M, some Γ in M and some assignment μ such that

$M, \Gamma, \mu \models \varphi \wedge \neg(\forall w)\psi(w)$, so

$$M, \Gamma, \mu \models \varphi \tag{2.25}$$

and

$$M, \Gamma, \mu \models (\exists w)\neg\psi(w),$$

i.e., for some $\hat{o} \in D(\Gamma)$,

$$M, \Gamma, \mu \models \neg\psi(w)[w : \hat{o}] \tag{2.26}$$

Let an assignment μ' be like μ except that $\mu'(z) = \hat{o}$, then since z is a variable that does not occur free in φ, and $\psi(z)$ is the result of replacing all free occurrences of w in $\psi(w)$ with occurrences of z , from (2.25) and (2.26),

$$M, \Gamma, \mu' \models \varphi \wedge \neg\psi(z),$$

hence $\varphi \supset \psi(z)$ is not valid.

R4 Suppose $\varphi(\vec{x})$ is valid, where $\vec{x}$ is $x_1, x_2, \ldots, x_n$, then for every Γ and every μ,

$\Gamma, \mu \models \varphi(\vec{x})$, so

$\Gamma, \mu \models \varphi[x_1 : o_1, \ldots, x_n : o_n]$, for every $o_1, \ldots, o_n$,

$\Gamma, \mu \models \varphi[x_1 : \mu(a_1, \Gamma), \ldots, x_n : \mu(a_n, \Gamma)]$,

$\Gamma, \mu \models \varphi(\vec{a})$, where $\vec{a}$ is $a_1, a_2, \ldots, a_n$ (since the $\vec{a}$ replacing $\vec{x}$ couldn't be Fregean occurrences).

■

2.8 Completeness

The completeness of the axiom system is proved in two steps: first, a canonical model for $\mathcal{L}_{\text{FMP}}$, which is not *normal*, is constructed; then we show that this canonical model can be "converted" into an *equivalent* normal model. From these the completeness result will follow. The following two definitions concern the two terms just mentioned.

Definition 2.8.1 A **FMP** model is *normal* if $\pi(=,\Gamma)$, the interpretation of "=" at a world Γ in that model, is the equality relation on the universe of that model: $\pi(=,\Gamma) = \{\langle o, o\rangle \mid o \in \bigcup_{\Delta\in G} D(\Delta)\}$ (where G is the set of all possible worlds of the model).

Definition 2.8.2 Two **FMP** models M and M' are *equivalent* if, for every closed formula φ, $M \models \varphi$ iff $M' \models \varphi$.

2.8.1 The Construction of a (Non-Normal) Canonical Model

$B-$Completeness and $B-$Saturatedness

Lemma 2.8.3 *Let $\theta = \neg B_i\langle x.\psi\rangle(a)$ (where ψ is a formula). If S is a consistent set of a sublanguage $\mathcal{L}$ of $\mathcal{L}_{\text{FMP}}$ containing θ, and z is a variable not in $\mathcal{L}$, then $S \cup \theta^{\mathcal{L}}(z)$ is consistent.*

Proof Suppose $S \cup \theta^{\mathcal{L}}(z)$ is inconsistent, then there is a finite subset S_0 of S, and a finite subset $\theta_0^{\mathcal{L}}(z)$ of $\theta^{\mathcal{L}}(z)$, such that $S_0 \cup \theta_0^{\mathcal{L}}(z)$ is inconsistent; this means

$$\vdash \bigwedge \theta_0^{\mathcal{L}}(z) \supset \neg \bigwedge S_0,$$

where $\bigwedge S_0$ is the conjunction of all formulas in S_0, and $\bigwedge \theta_0^{\mathcal{L}}(z)$ is the conjunction of all members of $\theta_0^{\mathcal{L}}(z)$. Since z is not in $\mathcal{L}$, Derived Rule **DR8** can be applied to obtain

$$\vdash \theta \supset \neg \bigwedge S_0,$$

hence $\theta \cup S_0$ is inconsistent. Consequently, S, with θ in it, is inconsistent. ■

Definition 2.8.4 Let S be a set of formulas and a be a constant in some sublanguage $\mathcal{L}$ of $\mathcal{L}_{\mathbf{FMP}}$. A term t is (a, i, S)-*good* w.r.t. $\mathcal{L}$ if

(1) $E(t) \in S$, and

(2) for every formula φ of $\mathcal{L}$, $B_i\langle x.\varphi\rangle(a) \supset B_i\varphi(t) \in S$.

Definition 2.8.5 Let $\theta = \neg B_i\langle x.\psi\rangle(a)$ and S be a set of formulas of some sublanguage $\mathcal{L}$ of $\mathcal{L}_{\mathbf{FMP}}$ containing θ. If there is some term t such that $\neg B_i\psi(t) \in S$, then we say that t is a *witness* to θ w.r.t. S. If, in addition, t is (a, i, S)-good w.r.t. $\mathcal{L}$, then we say that t is a *good witness* to θ w.r.t. S and $\mathcal{L}$.

Definition 2.8.6 [B—completeness] A set S of formulas of a sublanguage $\mathcal{L}$ of $\mathcal{L}_{\mathbf{FMP}}$ is *B-complete* w.r.t. $\mathcal{L}$ if every θ-formula in S has a good witness w.r.t. S and $\mathcal{L}$.

Definition 2.8.7 A set S of formulas of a first-order language $\mathcal{L}$ is *E-complete* in $\mathcal{L}$ if, for every formula φ of $\mathcal{L}$ and variable x, there is some term t in $\mathcal{L}$ such that $E(t) \wedge ((\exists x)\varphi \supset \varphi(t/x)) \in S$.

Definition 2.8.8 [B—saturatedness w.r.t. a language] A set S is *B-saturated* w.r.t. a sublanguage $\mathcal{L}$ of $\mathcal{L}_{\mathbf{FMP}}$ if it meets all the following three conditions:

(1) S is maximal consistent w.r.t. $\mathcal{L}$,

(2) S is E-complete in $\mathcal{L}$,

(3) S is B-complete w.r.t. $\mathcal{L}$.

Lemma 2.8.9 *Every consistent set S in a sublanguage $\mathcal{L}$ of $\mathcal{L}_{\mathbf{FMP}}$ can be extended to a* B-*saturated set w.r.t. an extension $\mathcal{L}^+$ of $\mathcal{L}$.*

Proof Let there be an infinite supply of variables not in $\mathcal{L}$. Let $\mathcal{L}^+$ be the language enlarged from $\mathcal{L}$ with these new variables. Arrange all formulas of $\mathcal{L}^+$ in some order: $\varphi_1, \varphi_2, \varphi_3, \ldots$.

Let $S_0 = S$.

For any positive integer n, if $S_{n-1} \cup \{\varphi_n\}$ is inconsistent, then let $S_n = S_{n-1}$; otherwise (if $S_{n-1} \cup \{\varphi_n\}$ is consistent), then:

a) If φ_n is not of the form $(\exists x)\psi$ or a θ-formula, then let

$$S_n = S_{n-1} \cup \{\varphi_n\};$$

b) If φ_n is $(\exists x)\psi(x)$, then let

$$S_n = S_{n-1} \cup \{(\exists x)\psi(x)\} \cup \{E(z)\} \cup \{\psi(z)\},$$

where z is a variable not occurring in S_{n-1};

c) If φ_n is θ, then let

$$S_n = S_{n-1} \cup \{\theta\} \cup \theta^{\mathcal{L}}(z),$$

where z is a variable not occurring in S_{n-1}.

Since there is an infinite supply of new variables, (b) and (c) can always be done, and by the usual argument about E-completeness and Lemma 2.8.3, each time (a), (b) or (c) is carried out, consistency is retained.

Clearly, $S = S_0 \subseteq S_1 \subseteq S_2 \subseteq \ldots$.

Finally, let

$$S^* = S_0 \cup S_1 \cup S_2 \cup S_3 \cup \ldots.$$

Then S^* is maximal consistent and E-complete in $\mathcal{L}^+$, by the usual argument about E-completeness. We claim that S^* is also B-complete w.r.t. $\mathcal{L}^+$. What remains is to show this claim.

Suppose $\theta \in S^*$, where $\theta = \neg B_i\langle x.\psi\rangle(a)$ (we need to show that θ has a good witness w.r.t. S^* and $\mathcal{L}^+$). Then $\theta = \varphi_k$ for some k, so by the construction of the S_n's,

$$\theta \in S_k$$

and

$$\theta^{\mathcal{L}}(z) \subseteq S_k \subseteq S^* \tag{2.27}$$

(where z is a variable not occurring in S_{k-1}).

By Definition 2.3.5, (2.27) implies

$$\bigwedge_{\varphi'}(\forall\vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(z)) \wedge E(z) \wedge \neg B_i\psi(z) \in S^*,$$

for the conjunction $\bigwedge_{\varphi'}$ of members φ' of any *finite* subset of the set of formulas of $\mathcal{L}$; since S^* is maximal consistent, this implies

$$\neg B_i\psi(z) \in S^*, \tag{2.28}$$

$$E(z) \in S^*, \tag{2.29}$$

and for every formula φ' of $\mathcal{L}$,

$$(\forall\vec{y})(B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(z)) \in S^* \tag{2.30}$$

where (2.30), in turn, implies that for every formula φ' of $\mathcal{L}$, and for every term $t_1, t_2, \ldots$,

$$B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(z) \in S^*, \tag{2.31}$$

where every occurrence of the element y_j of $\vec{y}$ in φ' is substituted by t_j.

Because all the formulas of $\mathcal{L}^+ - \mathcal{L}$ are a result of the introduction of the new variables, for every formula $\varphi(z_{i_1}, \ldots, z_{i_m})$ of $\mathcal{L}^+ - \mathcal{L}$, there is a corresponding formula $\varphi(y_{i_1}, \ldots, y_{i_m})$ of $\mathcal{L}$ which is one of the φ''s. By (2.31), when $y_{i_1}, \ldots, y_{i_m}$ in $\varphi(y_{i_1}, \ldots, y_{i_m})$ are substituted by $z_{i_1}, \ldots, z_{i_m}$, we must have

$$B_i\langle x.\varphi(z_{i_1}, \ldots, z_{i_m})\rangle(a) \supset B_i\varphi(z) \in S^*.$$

This means that not only when φ' is a formula of $\mathcal{L}$, but also when φ' is a formula of $\mathcal{L}^+ - \mathcal{L}$, the following is true:

$$B_i\langle x.\varphi'\rangle(a) \supset B_i\varphi'(z) \in S^*.$$

Consequently, for every formula φ of $\mathcal{L}^+$,

$$B_i\langle x.\varphi\rangle(a) \supset B_i\varphi(z) \in S^*. \tag{2.32}$$

By (2.28), z is a witness to θ w.r.t. S^*. By (2.29) and (2.32), z is (a, i, S^*)-good w.r.t. $\mathcal{L}^+$. Therefore z is a good witness to θ w.r.t. S^* and $\mathcal{L}^+$.

We conclude that every θ-formula in S^* has a good witness w.r.t. S^* and $\mathcal{L}^+$. So S^* is B-complete w.r.t. $\mathcal{L}^+$.

Therefore S^* is B-saturated w.r.t. $\mathcal{L}^+$. ■

Definition 2.8.10 [B–saturatedness] A set is *B-saturated* if it is B-saturated with respect to some language $\mathcal{L}$.

Definition 2.8.11 For any set S of formulas, $S^{-B_i} \overset{df}{=} \{\varphi \mid B_i\varphi \in S\}$.

Lemma 2.8.12 *If S is a B-saturated set, and $\neg B_i\psi \in S$, then $S^{-B_i} \cup \{\neg\psi\}$ is consistent.*

Proof Follow the same line of reasoning as in, *e.g.*, [19]. ■

The Canonical Model M^*

We now construct the model

$$M^* = \langle G^*, I, R_1^*, \ldots, R_{|I|}^*, D^*, \sigma^*, m^*, \pi^* \rangle$$

in the following way:

1. G^* is the collection of all B-saturated sets of formulas (for each $\Gamma \in G^*$, $\mathcal{L}_-$ denotes the smallest language in which the formulas in Γ are written);

2. For any $i \in I$ and any $\Gamma, \Delta \in G^*$, $\Gamma R_i^* \Delta$ iff $\Gamma^{-B_i} \subseteq \Delta$;

3. For any $\Gamma \in G^*$, $D^*(\Gamma) = \{t \mid t$ is a term in $\mathcal{L}_-$ and $E(t) \in \Gamma\}$;

4. For each $a \in C$, $\Gamma \in G^*$, $\sigma^*(a, \Gamma) = a$;

5. For any $a \in C$, $i \in I, \Gamma \in G^*$,

   $m^*(a, i, \Gamma) = \{t \mid t$ is a term in $\mathcal{L}_-$ and t is (a, i, Γ)-good w. r. t. $\mathcal{L}_-\}$;

6. $\langle t_1, \ldots, t_n \rangle \in \pi^*(P, \Gamma)$ iff $P(t_1, \ldots, t_n) \in \Gamma$ (where $P \in \mathcal{P}^n$);

7. $\langle t_1, t_2\rangle \in \pi^*(=,\Gamma)$ iff $t_1 = t_2 \in \Gamma$

Finally, let ρ be an assignment such that for each variable x, $\rho(x) = x$.

Proposition 2.8.13 *(1) For every term t and every $\Gamma \in G^*$, $\rho(t,\Gamma) = t$.*

(2) For any $\Gamma \in G^*$ and any instance of $\neg B_i\langle x.\psi\rangle(a) \in \Gamma$, if t is a good witness to the latter w.r.t. Γ and $\mathcal{L}_-$, then $t \in m^*(a,i,\Gamma)$.

Lemma 2.8.14 *M^* is a* **FMP** *model (except that it is non-normal).*

Proof 1. R_i^* is serial: follows from **FMP b3**.

2. R_i^* is transitive: follows from **FMP b4**.

3. R_i^* is Euclidean: follows from **FMP b5**.

4. D^* is monotonic:

   Suppose $t \in D^*(\Gamma)$, then t is a term in $\mathcal{L}_-^+$ and $E(t) \in \Gamma$;

   by **FMP 2** and **R3**, $B_i(\forall x)E(x) \in \Gamma$, so by Converse Barcan Formula (Lemma 2.6.8), $(\forall x)B_iE(x) \in \Gamma$. Thus

   $(\forall x)B_iE(x) \wedge E(t) \in \Gamma$, hence by **FMP 3**,

   $B_iE(t) \in \Gamma$, therefore

   $E(t) \in \Gamma^{-B_i}$, so

   $E(t) \in \Delta$, for every Δ such that $\Gamma R_i^* \Delta$, hence

   $t \in D^*(\Delta)$, for every Δ such that $\Gamma R_i^* \Delta$.

   Therefore, for any $\Gamma, \Delta \in G^*$, if $\Gamma R_i^* \Delta$, then $D^*(\Gamma) \subseteq D^*(\Delta)$.

5. σ^* meets the local rigidity condition—in fact, it meets the *global rigidity condition:*

   For any a, any $\Gamma, \Delta \in G^*$ (regardless whether $\Gamma R_i \Delta$ for some i or not),

   $\sigma^*(a,\Gamma) = a = \sigma^*(a,\Delta)$.

6. For any $a \in C$, $i \in I$, $\Gamma \in G^*$, $m^*(a,i,\Gamma) \subseteq D^*(\Gamma)$: follows from the definition of m^*.

■

Lemma 2.8.15 *For any* $i \in I$, $a \in C$, $\Gamma \in G^*$, *and any pseudo-formula* φ,

$B_i\langle x.\varphi\rangle(a) \in \Gamma$ *iff for every term* $t \in m^*(a,i,\Gamma)$, $B_i\varphi(t) \in \Gamma$.

Proof We prove it by induction on the nesting complexity (cf. Definition 2.3.3) of φ.

(I) Base: φ is of nesting complexity 0, *i.e.*, φ is a formula:

($\Longrightarrow$)

Suppose $B_i\langle x.\varphi\rangle(a) \in \Gamma$.

Let $t \in m^*(a,i,\Gamma)$. Then by the definition of m^*, t is (a,i,Γ)-good w.r.t. $\mathcal{L}_-$, hence, by Definition 2.8.4, $B_i\langle x.\varphi\rangle(a) \supset B_i\varphi(t) \in \Gamma$. Since $B_i\langle x.\varphi\rangle(a) \in \Gamma$, and Γ is maximal consistent, it follows that $B_i\varphi(t) \in \Gamma$. Hence for every $t \in m^*(a,i,\Gamma)$, $B_i\varphi(t) \in \Gamma$.

($\Longleftarrow$)

Suppose $B_i\langle x.\varphi\rangle(a) \notin \Gamma$, then $\neg B_i\langle x.\varphi\rangle(a) \in \Gamma$. Since Γ is $\mathcal{B}$-saturated, $\neg B_i\langle x.\varphi\rangle(a)$ has a good witness t_0 w.r.t. Γ and $\mathcal{L}_-$ such that $\neg B_i\varphi(t_0) \in \Gamma$, *i.e.*, $B_i\varphi(t_0) \notin \Gamma$. Also, by Part (2) of Proposition 2.8.13, $t_0 \in m^*(a,i,\Gamma)$. Hence for some t (namely t_0) such that $t \in m^*(a,i,\Gamma)$, $B_i\varphi(t) \notin \Gamma$.

(II) Induction: φ is $\langle y.\psi\rangle(b)$, where ψ is a pseudo-formula of nesting complexity k:

By Definition 2.3.2 and Axiom **FMP bs1**,

$B_i\langle x.\langle y.\psi\rangle(b)\rangle(a) \in \Gamma$

iff $(\forall x)\{(\forall y)[B_i\langle x.\psi\rangle(a) \supset B_i\psi] \supset B_i\langle y.\psi\rangle(b)\} \in \Gamma$

iff for every term $t \in D^*(\Gamma)$, either

$$\neg(\forall y)[B_i\langle x.\psi\rangle(a) \supset B_i\psi(t,y)] \in \Gamma \tag{2.33}$$

or

$$B_i\langle y.\psi(t,y)\rangle(b) \in \Gamma. \tag{2.34}$$

By $E-$completeness, (2.33) means that there is some term $\hat{t}$ such that $E(\hat{t}) \in \Gamma$ and

$$B_i\langle x.\psi(x,\hat{t})\rangle(a) \in \Gamma \tag{2.35}$$

and

$$\neg B_i\psi(t,\hat{t}) \in \Gamma. \tag{2.36}$$

By the inductive hypothesis, (2.35) is true iff for every $t_1 \in m^*(a,i,\Gamma)$,

$$B_i\psi(t_1,\hat{t}) \in \Gamma. \tag{2.37}$$

Now, in (2.36), since t is any member of the domain $D^*(\Gamma)$, and $m^*(a,i,\Gamma) \subseteq D^*(\Gamma)$, we thus have: for every $t_1 \in m^*(a,i,\Gamma)$,

$$\neg B_i\psi(t_1,\hat{t}) \in \Gamma, \tag{2.38}$$

which contradicts (2.37).

Therefore, if t is any $t_1 \in m^*(a,i,\Gamma)$, (2.37) and (2.38) cannot be both true, *i.e.*, (2.33) cannot be true; hence (2.34) must instead be true: for every $t_1 \in m^*(a,i,\Gamma)$,

$$B_i\langle y.\psi(t_1,y)\rangle(b) \in \Gamma,$$

i.e., $B_i\varphi(t_1) \in \Gamma$. ∎

Corollary 2.8.16 *For any* $i \in I$, $a_1,\ldots,a_n \in C$, $\Gamma \in G^*$, *and any formula* φ, $B_i\langle x_1 \ldots x_n.\varphi\rangle(a_1,\ldots,a_n$ Γ *iff for every* t_j *such that* $t_j \in m^*(a_j,i,\Gamma)$ $(j = 1,\ldots,n)$, $B_i\varphi(t_1,\ldots,t_n) \in \Gamma$.

Lemma 2.8.17 (Main lemma) *For any* $\Gamma \in G^*$, *any formula* φ *in* $\mathcal{L}_{\mathbf{FMP}}$,

$$M^*,\Gamma,\rho \models \varphi \Longleftrightarrow \varphi \in \Gamma.$$

Proof By induction on the complexity of φ.

1. φ is atomic:

   $M^*,\Gamma,\rho \models P(t_1,\ldots,t_n)$ (where $P \in \mathcal{P} - \{E\}$)

iff $\langle \rho(t_1,\Gamma),\ldots,\rho(t_n,\Gamma)\rangle \in \pi^*(P,\Gamma)$

iff $\langle t_1,\ldots,t_n\rangle \in \pi^*(P,\Gamma)$

iff $P(t_1,\ldots,t_n) \in \Gamma$.

$M^*,\Gamma,\rho \models t_1 = t_2$

iff $\langle \rho(t_1,\Gamma),\rho(t_2,\Gamma)\rangle \in \pi^*(=,\Gamma)$

iff $\langle t_1,t_2\rangle \in \pi^*(=,\Gamma)$

iff $t_1 = t_2 \in \Gamma$.

$M^*,\Gamma,\rho \models E(t)$

iff $\rho(t,\Gamma) \in D^*(\Gamma)$

iff $t \in D^*(\Gamma)$

iff $E(t) \in \Gamma$.

2. $\varphi = \neg\psi$:

   $M^*,\Gamma,\rho \models \neg\psi$

   iff $M^*,\Gamma,\rho \not\models \psi$

   iff $\psi \notin \Gamma$ (inductive hypothesis)

   iff $\neg\psi \in \Gamma$ (maximal consistency of Γ).

3. $\varphi = \psi_1 \wedge \psi_2$:

   $M^*,\Gamma,\rho \models \psi_1 \wedge \psi_2$

   iff $M^*,\Gamma,\rho \models \psi_1$ and $M^*,\Gamma,\rho \models \psi_2$

   iff $\psi_1 \in \Gamma$ and $\psi_2 \in \Gamma$ (inductive hypothesis)

   iff $\psi_1 \wedge \psi_2 \in \Gamma$ (maximal consitency of Γ).

4. $\varphi = (\exists x)\psi(x)$:

   $M^*, \Gamma, \rho \models (\exists x)\psi(x)$

   iff $M^*\Gamma, \rho \models \psi(x)[x : o]$ for some $o \in D^*(\Gamma)$

   iff $M^*, \Gamma, \rho \models \psi(x)[x : t]$ for some $t \in D^*(\Gamma)$

   iff $M^*, \Gamma, \rho \models \psi(x)[x : t]$ for some term t such that $E(t) \in \Gamma$

   iff $M^*, \Gamma, \rho \models \psi(t)$ for some term t such that $E(t) \in \Gamma$ (proof details omitted)

   iff $\psi(t) \in \Gamma$ for some term t such that $E(t) \in \Gamma$ (inductive hypothesis)

   iff $(\exists x)\psi(x) \in \Gamma$ (E-completeness for the 'if' part).

5. $\varphi = B_i\psi$, where $\psi \in \Phi$:

   (a) Suppose $B_i\psi \in \Gamma$, then $\psi \in \Gamma^{-B_i}$, so for every Δ such that $\Gamma^{-B_i} \subseteq \Delta$,

   $\psi \in \Delta$; hence for every Δ such that $\Gamma R_i^* \Delta$,

   $M^*, \Delta, \rho \models \psi$ (by the inductive hypothesis), hence

   $M^*, \Gamma, \rho \models B_i\psi$ (by **B-r**).

   (b) Conversely, suppose $B_i\psi \notin \Gamma$, then

   $\neg B_i\psi \in \Gamma$ (since Γ is maximal consistent),

   so by Lemma 2.8.12, $\Gamma^{-B_i} \cup \{\neg\psi\}$ is consistent; therefore, by Lemma 2.8.9, there is some $\Delta \in G^*$ such that $\Gamma^{-B_i} \cup \{\neg\psi\} \subseteq \Delta$, so

   $\Gamma^{-B_i} \subseteq \Delta$ and $\neg\psi \in \Delta$;

   from the former, $\Gamma R_i^* \Delta$; from the latter, $M^*, \Delta, \rho \models \neg\psi$ (by the inductive hypothesis). Therefore, there is some Δ such that $\Gamma R_i^* \Delta$ and

   $M^*, \Delta, \rho \models \neg\psi$, thus

   $M^*, \Gamma, \rho \models \neg B_i\psi$.

6. $\varphi = B_i\langle x.\psi\rangle(a)$, where ψ is a pseudo-formula.

   $M^*, \Gamma, \rho \models B_i\langle x.\psi\rangle(a)$

iff for every $o \in m^*(a, i, \Gamma)$, $M^*, \Gamma, \rho \models B_i\psi[x : o]$ (by **B-d**)

iff for every $t \in m^*(a, i, \Gamma)$, $M^*, \Gamma, \rho \models B_i\psi[x : t]$

iff for every $t \in m^*(a, i, \Gamma)$, $M^*, \Gamma, \rho \models B_i\psi(t)$ (provable by induction on the complexity of ψ)

iff for every t such that $t \in m^*(a, i, \Gamma)$, $B_i\psi(t) \in \Gamma$ (by the inductive hypothesis).

■

2.8.2 Converting M^* to a Normal Model

Converting M^* to a Special Model

Definition 2.8.18 Given a **FMP** model M (not necessarily normal) and any two possible worlds Γ, Δ of M, we say Γ and Δ are *directly connected* if $\Gamma R_i \Delta$ for some i, or $\Delta R_i \Gamma$ for some i. Γ and Δ are *connected* if there are $\Sigma_1, \ldots \Sigma_k (k \geq 1)$ such that Σ_1 is Γ, Σ_k is Δ, and for every j such that $1 \leq j < k$, Σ_j and Σ_{j+1} are directly connected.

Proposition 2.8.19 *For a given* **FMP** *model M, connectedness is an equivalence relation on the set of possible worlds.*

Definition 2.8.20 [special model] A **FMP** model (not necessarily normal) $M = \langle G, I, R_1, \ldots, R_{|I|}, D, \sigma, m, \pi \rangle$ is a *special model* if for every $\Gamma, \Delta \in G$ and every o, o' in the universe of M, $\langle o, o' \rangle \in \pi(=, \Gamma)$ iff $\langle o, o' \rangle \in \pi(=, \Delta)$.

Lemma 2.8.21 *Every model M in which the Axioms* **FMP b6a** *and* **FMP b6b** *are valid has an equivalent special model M'; in particular, there is a special* **FMP** *model equivalent to M^*.*

Proof First we claim that the Axioms **FMP b6a** and **FMP b6b** are valid in M^*. This follows from (i) the definition of M^*, in which the worlds are maximal consistent sets of formulas of $\mathcal{L}_{\mathbf{FMP}}$, and (ii) Lemma 2.8.17.

Now let $M = \langle G, I, R_1, \ldots, R_{|I|}, D, \sigma, m, \pi\rangle$ be any model in which the Axioms **FMP b6a** and **FMP b6b** are valid, so for any i,

$$M \models x = y \supset B_i(x = y)$$

and

$$M \models x \neq y \supset B_i(x \neq y).$$

Thus, for every world Γ of M, and every o, o', if $\Gamma \models x = y[x : o; y : o']$, then for every Δ such that $\Gamma R_i \Delta$, $\Delta \models x = y[x : o; y : o']$; that is,

for every world Γ of M, and every o, o', if $\langle o, o'\rangle \in \pi(=, \Gamma)$, then for every Δ such that $\Gamma R_i \Delta$, $\langle o, o'\rangle \in \pi(=, \Delta)$.

Similarly, for every world Γ of M, and every o, o', if $\langle o, o'\rangle \notin \pi(=, \Gamma)$, then for every Δ such that $\Gamma R_i \Delta$, $\langle o, o'\rangle \notin \pi(=, \Delta)$.

So for every Γ and every Δ such that $\Gamma R_i \Delta$, and for every o, o',

$\langle o, o'\rangle \in \pi(=, \Gamma)$ iff $\langle o, o'\rangle \in \pi(=, \Delta)$.

Hence it follows, by Definition 2.8.18, that for any two connected worlds Γ, Δ, and for every o, o',

$$\langle o, o'\rangle \in \pi(=, \Gamma) \text{ iff } \langle o, o'\rangle \in \pi(=, \Delta). \tag{2.39}$$

To show that for *every* Γ, Δ, connected or not, (2.39) holds for every o, o', we need only be reminded that the canonical model M^* is meant to be a counter model for any formula not provable in **FMP**, that is, for any given formula φ which is not provable in **FMP**, there is a world, say Σ, in M^* at which the formula is false. Since the worlds which are not connected with Σ are irrelevant to φ in this respect, they need not be included in the counter model for φ. Discarding these worlds, the result is a model in which all worlds are connected with one another. It then follows that for *every* Γ, Δ, (2.39) holds for every o, o'.

Thus, for any given formula, there is a special model which is equivalent to the original model M (as far as the formula is concerned). ■

Converting a Special Model to a Normal Model

Definition 2.8.22 Given a special model $M = \langle G, I, R_1, \ldots, R_{|I|}, D, \sigma, m, \pi \rangle$, we write $o \approx_M o'$ if, for some $\Gamma \in G$, $\langle o, o' \rangle \in \pi(=, \Gamma)$.

Proposition 2.8.23 *For any special model $M = \langle G, I, R_1, \ldots, R_{|I|}, D, \sigma, m, \pi \rangle$,*

(1) $\approx_M$ is an equivalence relation.

(2) If $o_1, \ldots, o_n, o'_1, \ldots, o'_n \in D(\Gamma)$, and $o_1 \approx_M o'_1, \ldots, o_n \approx_M o'_n$, then for any $P \in \mathcal{P}_\backslash$, $\langle o_1, \ldots, o_n \rangle \in \pi(P, \Gamma)$ iff $\langle o'_1, \ldots, o'_n \rangle \in \pi(P, \Gamma)$.

Proof Follows from the fact that any interpretation of $=$, $\pi(=, \Gamma)$, must meet reflexivity and the Substitution Principle of "=", and the fact that any two unconnected worlds have disjoint domains. ∎

Lemma 2.8.24 *(1) For every non-normal special model M, there is an equivalent normal model $\overline{M}$.*

(2) For every non-normal special **FMP** *model M, there is an equivalent normal* **FMP** *model $\overline{M}$.*

Proof We prove Part (2) only, which is sufficient for Part (1).

Let $M = \langle G, I, R_1, \ldots, R_{|I|}, D, \sigma, m, \pi \rangle$ be a (non-normal) special **FMP** model.

By Proposition 2.8.23, $\approx_M$ partitions the universe of M into equivalence classes:

$[o] \overset{df}{=} \{o' \mid o \approx_M o'\}$.

Let $\overline{o} \overset{df}{=} [o]$.

Define a new model $\overline{M}$ as follows:

$$\overline{M} = \langle G, I, R_1, \ldots, R_{|I|}, \overline{D}, \overline{\sigma}, \overline{m}, \overline{\pi} \rangle,$$

where, for any $\Gamma \in G$, any a, i,

$\overline{D}(\Gamma) = \{\overline{o} \mid o' \in D(\Gamma) \text{ for some } o' \in \overline{o}\}$,

$\overline{\sigma}(a, \Gamma) = \overline{\sigma(a, \Gamma)}$,

$\overline{m}(a, i, \Gamma) = \{\overline{o} \mid o' \in m(a, i, \Gamma), \text{for some } o' \in \overline{o}\}$,

$\langle \overline{o}_1, \ldots, \overline{o}_n \rangle \in \overline{\pi}(P, \Gamma)$ iff for some $o_1, \ldots, o_n \in D(\Gamma)$ such that $o_1 \in \overline{o}_1, \ldots, o_n \in \overline{o}_n$, $\langle o_1, \ldots, o_n \rangle \in \pi(P, \Gamma)$.

Claim 1 For any a, i and any $\Gamma \in G$, $\overline{m}(a, i, \Gamma) \subseteq \overline{D}(\Gamma)$:

Suppose $\overline{o} \in \overline{m}(a, i, \Gamma)$, then for some $o' \in \overline{o}$, $o' \in m(a, i, \Gamma) \subseteq D(\Gamma)$, therefore $\overline{o} \in \overline{D}(\Gamma)$.

Claim 2 $\overline{D}$ is monotonic:

Suppose $\Gamma R_i \Delta$ and $\overline{o} \in \overline{D}(\Gamma)$, then for some $o' \in \overline{o}$, $o' \in D(\Gamma)$. Since M is a **FMP** model (being a special model), D is monotonic, that is, $D(\Gamma) \subseteq D(\Delta)$, so it follows from that for some $o' \in \overline{o}$, $o' \in D(\Delta)$, thus $\overline{o} \in \overline{D}(\Delta)$. Therefore $\overline{D}(\Gamma) \subseteq \overline{D}(\Delta)$, *i.e.*, $\overline{D}$ is monotonic.

Claim 3 $\overline{\sigma}$ meets the local rigidity condition:

$\overline{\sigma}(a, \Gamma) = \overline{\sigma(a, \Gamma)} = [\sigma(a, \Gamma)] = \{o \mid \sigma(a, \Gamma) \approx_M o\}$

$= \{o \mid \text{for some } \Omega \in G, \langle \sigma(a, \Gamma), o \rangle \in \pi(=, \Omega)\}$;

since M is a **FMP** model, for any Δ such that $\Gamma R_i \Delta$ for some i,

$\sigma(a, \Gamma) = \sigma(a, \Delta)$ (local rigidity);

hence, by the above,

$\overline{\sigma}(a, \Gamma) = \overline{\sigma}(a, \Delta)$.

Claim 4 $\overline{M}$ is a normal model:

$\overline{M}, \Gamma \models \overline{o}_1 = \overline{o}_2$

iff $\langle \overline{o}_1, \overline{o}_2 \rangle \in \overline{\pi}(=, \Gamma)$

iff for some $o_1, o_2 \in D(\Gamma)$ such that $o_1 \in \overline{o}_1$ and $o_2 \in \overline{o}_2$, $\langle o_1, o_2 \rangle \in \pi(=, \Gamma)$,

iff for some $o_1, o_2 \in D(\Gamma)$ such that $o_1 \in \overline{o}_1$ and $o_2 \in \overline{o}_2$, $o_1 \approx_M o_2$

iff $\overline{o}_1, \overline{o}_2 \in \overline{D}(\Gamma)$ and $\overline{o}_1 = \overline{o}_2$.

Therefore $\overline{\pi}(=, \Gamma) = \{\langle o, o \rangle \mid o \in \overline{D}(\Gamma)\}$, so $\overline{M}$ is normal.

Claim 5 For any formula φ, $M, \Gamma, \mu \models \varphi$ iff $\overline{M}, \Gamma, \overline{\mu} \models \varphi$,

where $\overline{\mu}$ is an assignment such that $\overline{\mu}(x) = \overline{\mu(x)}$.

This is shown by induction on the complexity of φ:

i) φ is $P \in \mathcal{P}$ (including E and $=$):

$\overline{M}, \Gamma, \overline{\mu} \models P(t_1, \ldots, t_n)$

iff $\langle \overline{\mu}(t_1, \Gamma), \ldots, \overline{\mu}(t_n, \Gamma) \rangle \in \overline{\pi}(P, \Gamma)$

iff $\langle \overline{\mu(t_1, \Gamma)}, \ldots, \overline{\mu(t_n, \Gamma)} \rangle \in \overline{\pi}(P, \Gamma)$

iff for some $o_1, \ldots, o_n \in D(\Gamma)$ such that $o_1 \in \overline{\mu(t_1, \Gamma)}, \ldots, o_n \in \overline{\mu(t_n, \Gamma)}$, $\langle o_1, \ldots, o_n \rangle \in \pi(P, \Gamma)$

iff for some $o_1, \ldots, o_n \in D(\Gamma)$ such that $o_1 \approx_M \mu(t_1, \Gamma), \ldots, o_n \approx_M \mu(t_n, \Gamma)$, $\langle o_1, \ldots, o_n \rangle \in \pi(P, \Gamma)$

iff $\langle \mu(t_1, \Gamma), \ldots, \mu(t_n, \Gamma) \rangle \in \pi(P, \Gamma)$ (by Part (2) of Proposition 2.8.23)

iff $M, \Gamma, \mu \models P(t_1, \ldots, t_n)$.

ii) φ is $\psi_1 \wedge \psi_2$ or $\neg\psi$: straightforward using the inductive hypothesis.

iii) φ is $(\exists x)\psi(x)$:

$\overline{M}, \Gamma, \overline{\mu} \models (\exists x)\psi(x)$

iff for some $\overline{o} \in \overline{D}(\Gamma)$, $\overline{M}, \Gamma, \overline{\mu} \models \psi(x)[x : \overline{o}]$

iff for some $o \in D(\Gamma)$, $M, \Gamma, \mu \models \psi(x)[x : o]$ (by the definition of $\overline{D}$ and the inductive hypothesis)

iff $M, \Gamma, \mu \models (\exists x)\psi(x)$.

iv) φ is $B_i\psi$: straightforward using the inductive hypothesis.

v) φ is $B_i\langle x.\psi\rangle(a)$, where ψ is a pseudo-formula:

$\overline{M}, \Gamma, \overline{\mu} \models B_i\langle x.\psi\rangle(a)$

iff for every $\overline{o} \in \overline{m}(a, i, \Gamma)$, $\overline{M}, \Gamma, \overline{\mu} \models B_i\psi[x : \overline{o}]$ (by **B-d**) iff for every $o \in m(a, i, \Gamma) \subseteq D(\Gamma)$, there is some $o' \approx_M o$ (*i. e.*, $o' \in \overline{o}$) such that $M, \Gamma, \mu \models B_i\psi[x : o']$ (by the definition of $\overline{m}$ and the inductive hypothesis)

iff for every $o \in m(a, i, \Gamma)$, $M, \Gamma, \mu \models B_i\psi[x : o]$ (by Part (2) of Proposition 2.8.23)

iff $M, \Gamma, \mu \models B_i\langle x.\psi\rangle(a)$.

By Claims 1, 2 and 3, $\overline{M}$ is a **FMP** model. By Claim 4, $\overline{M}$ is a normal model. By Claim 5, $\overline{M}$ is equivalent to M. ■

2.8.3 The Completeness Result

Theorem 2.8.25 (Completeness) *The axiom system of* **FMP** *is complete with respect to the* **FMP** *semantics.*

Proof Suppose φ is consistent, then $\varphi \in \Gamma$ for some $\Gamma \in G^*$, so by Lemma 2.8.17, $M^*, \Gamma, \rho \models \varphi$ for some $\Gamma \in G^*$.

By Lemma 2.8.21, there is a special **FMP** model $M^{*\prime}$ and an assignment ρ' such that $M^{*\prime}, \Gamma, \rho' \models \varphi$ for some $\Gamma \in G^*$ of $M^{*\prime}$;

by Lemma 2.8.24, there is a normal **FMP** model $\overline{M^{*\prime}}$ and an assignment $\overline{\rho'}$ such that $\overline{M^{*\prime}}, \Gamma, \overline{\rho'} \models \varphi$ for some $\Gamma \in G^*$ of $\overline{M^{*\prime}}$.

Therefore φ is satisfiable in a normal **FMP** model.

Consequently, every consistent formula is satisfiable in a (normal) **FMP** model. ■

Chapter 3

Applications to Philosophical Issues

3.1 Some Aspects of FMP

3.1.1 Each Name is Believed to Refer to a Unique Object

We observe that though a mode of presentation $m(a, i, \Gamma)$ may have more than one member, it does not follow that i *believes* there is more than one thing called 'a'. In fact, we have:

Theorem 3.1.1 *For any i and any constant a,*

$$\vdash B_i \langle x.(\exists y)(\forall z)(z = x \equiv z = y) \rangle (a)$$

(for any believer i and any name 'a', i believes that 'a' names a unique thing).

Proof It's validity (hence provable) is straightforward, using the monotonicity of the domains and the fact that for any world Γ, $m(a, i, \Gamma) \subseteq D(\Gamma)$. ∎

3.1.2 Mode of Presentations Which Represent Contradictions

Suppose that someone i has a contradictory mode of presentation for the name 'a' *e.g.*, 'being round and being square' what would be the consequence? It seems natural to suppose that he would believe anything about a (as he understands the name): he would believe that a is round, that a is not round, a exists and a does not exist, etc. This is indeed reflected by our semantics: since $m(a, i, \Gamma) = \emptyset$ in such a case, by **B-d**,

$\Gamma \models B_i\langle x.\varphi\rangle(a)$ iff for every $o \in m(a, i, \Gamma) = \emptyset$, $\Gamma \models B_i\varphi[x : o]$,

and so regardless what φ is, $B_i\langle x.\varphi\rangle(a)$ would always be true.

3.1.3 Identities

De Re and *De Dicto* Beliefs about Identities

1. $B_i\langle xy.x = y\rangle(a, b)$ expresses "i believes that a is b."

   By **B-d**, this is true at a world Γ iff either both $m(a, i, \Gamma)$ and $m(b, i, \Gamma)$ are $\emptyset$, or they are both singletons and have the same member, that is, either both modes of presentation of a and b for i at Γ are of contradictory sort, or they are met uniquely by the same object. This is a *de dicto* belief about both 'a' and 'b'.

2. $B_i\langle x.x = b\rangle(a)$ expresses "of the thing named b, i believes that a is that thing."

   By **B-d**, this is true at a world Γ iff $m(a, i, \Gamma) = \{\sigma(b, \Gamma)\}(= \{\sigma(b, \Delta)\}$ for any Δ such that $\Gamma R_i \Delta$); that is, iff the thing designated by b is uniquely satisfies the mode of presentation of 'a' for i at Γ. In this case, the belief is *de dicto* about 'a' but *de re* about 'b'.

Remark According to the semantics **B-d**, $B_i\langle xy.x \neq y\rangle(a, b)$ is true at Γ iff $m(a, i, \Gamma) \cap m(b, i, \Gamma) = \emptyset$; that is, if someone believes that a is not b, then there must be no object which satisfies *both* the mode of presentation of a for him and that of b for him. This may give the impression that the semantics is somewhat counter-intuitive, as in real life it's possible, for example, that a person believes that Smith is not Jones, when the mode of presentation of 'Smith' for him is 'being the tallest man in the world', and that of 'Jones' for him is 'being the fattest man in the world', while there *is* someone who is both the tallest and the fattest man in the world.

Some Theorems About Identities

Theorem 3.1.2 *The following are theorems of* **FMP***:*

1. $\vdash B_i\langle x.x = x\rangle(a)$.
2. $\vdash B_i\langle x.\varphi\rangle(a) \wedge B_i\langle x.x = b\rangle(a) \supset B_i\varphi(b)$.
3. $\vdash B_i\langle x.\varphi\rangle(a) \wedge B_i\langle xy.x = y\rangle(a, b) \supset B_i\langle x.\varphi\rangle(b)$.
4. $B_i\langle x.\varphi\rangle(a) \wedge B_i\langle x.\neg\varphi\rangle(b) \supset B_i\langle xy.x \neq y\rangle(a, b)$, *where x has at least one free occurrence in φ.*

In the example about Lois and Superman/Clark Kent, the above items translate into:

1. Lois believes that Superman is Superman. (And: Lois believes that Clark Kent is Clark Kent)
2. If Lois believes that Superman (as she understands the name) flies, and what Lois thinks to be Superman is actually Clark Kent, then, of Clark Kent, Lois believes that he flies.
3. If Lois believes that Superman flies, and Lois believes that Superman is Clark Kent, then she believes that Clark Kent flies (as she understands the two names).
4. If Lois believes that Superman flies, and she believes that Clark Kent does not fly, then she must believe that Superman is not Clark Kent (as she understands the two names).

Examples

Example 3.1.3 Suppose whenever Peter hears or utters the name ‘Phosphorus’, he thinks of the large bright heavenly body in the sky he sees very often at night among many stars, which in fact is the moon; that is, the mode of presentation of the name ‘Phosphorus’ for Peter is “being the large bright heavenly body which I see very often in the sky at night”. Suppose now that Peter says:

'Phosphorus is a planet of the earth',

what he actually means is that the moon (which he thought is called 'Phosphorus') is a planet of the earth. That is, *of* the moon, Peter believes that it is a planet of the earth (a true belief). We can infer this *de re* belief of Peter because the moon is what satisfies the mode of presentation of 'Phosphorus' for Peter. Now we know that Phosphorus is Hesperus, but we cannot infer from Peter's utterance above that of Hesperus, Peter believes that it is a planet of the earth, because the name 'Phosphorus' in his utterance has the sense *he* (Peter) associates with it, not the sense *we* associate with the name 'Phosphorus'.

Example 3.1.4 [Version a] Suppose Lois says: "Superman can fly. Clark Kent cannot fly." We report it as:

> Lois believes that Superman (as she understands the name) can fly, and she does not believe that Clark Kent (as she understands the name) can fly.

The formal representation of this report is:

$$B_l\langle x.F\rangle(s) \wedge \neg B_l\langle x.F\rangle(k) \tag{3.1}$$

If the mode of presentation of 'Superman' for Lois is some description which Superman, the person actually designated by this name (who is also called Clark Kent) actually fits, then, if we use $\hat{o}$ to represent the person Superman (Clark Kent), and Γ to denote the real world, the following is true for the model which would describe the situation:

$$\hat{o} \in m(s, l, \Gamma) \tag{3.2}$$

$m(s, l, \Gamma)$ may or may not contain additional members: $m(s, l, \Gamma) = \{\hat{o}, \ldots\}$. The first conjunct in (3.1) would be true iff for every member o of $m(s, l, \Gamma)$, $B_l F[x : o]$ is true at Γ, which implies that $B_l F[x : \hat{o}]$ must be true, that is, *of* Superman, Lois believes that he can fly; but since Superman *Clark Kent* ($\sigma(s, \Gamma) = \sigma(k, \Gamma) = \hat{o}$, this also means that *of* Clark Kent, Lois believes that he can fly. On the other hand, the second conjunct of (3.1) could be true iff for some $\bar{o} \in m(k, l, \Gamma)$, $B_l F[x : \bar{o}]$ is false at Γ. Therefore, the model must be such that there is at least one member ($\bar{o}$) in $m(k, l, \Gamma)$ which is not in $m(s, l, \Gamma)$ (given that (3.1) is the case). As for the real Superman (Clark Kent), $\hat{o}$, it may or may not be in $m(k, l, \Gamma)$: for example, suppose the mode of presentation of Clark Kent for Lois is indeed satisfied by Clark Kent, *i.e.*, $\hat{o} \in m(k, l, \Gamma)$, so $m(k, l, \Gamma) = \{\hat{o}, \bar{o}, \ldots\}$. In this case, (3.1) holds, and the following *de re* belief report holds without a contradiction:

$$B_l F(s) \tag{3.3}$$

which is equivalent to

$$B_l F(k), \tag{3.4}$$

that is,

Of Superman (Clark Kent), Lois believes that he can fly.

If Lois has made some mistake in her understanding of the name 'Superman' so that the mode of presentation of 'Superman' for her is in fact not satisfied by Superman (Clark Kent) at all, then $\hat{o} \notin m(s, l, \Gamma)$, and a model can still be constructed to describe the situation; no contradiction, *de re* or *de dicto*, would arise. Further detailed illustrations are omitted.

Example 3.1.5 [Version b] Suppose, as before, Lois says: "Superman can fly. Clark Kent cannot fly." But this time we report it as:

Lois believes that Superman can fly and that Clark Kent can not fly.

This report is stronger than the previous one.The formal representation of it is:

$$B_l\langle x.F\rangle(s) \wedge B_l\langle x.\neg F\rangle(k) \tag{3.5}$$

Suppose again that the mode of presentation of 'Superman' for Lois is some description which Superman (Clark Kent) indeed fits, then, as before, using $\hat{o}$ to represent the person Superman (Clark Kent), and Γ the real world, (3.2) should be true for the model for such a case, that is, $m(s, l, \Gamma) = \{\hat{o}, \ldots\}$. For the first conjunct in (3.5) to be true, it must be that for every member o of $m(s,l,\Gamma)$, $B_lF[x:o]$ is true at Γ, so $B_lF[x:o]$ is true for all $o \in m(s,l,\Gamma)$, in particular,

$$B_lF[x:\hat{o}] \tag{3.6}$$

is true, *i.e.*,

Of Superman (Clark Kent), Lois believes that he can fly.

Now, what would be the mode of presentation of Clark Kent for Lois in this case? If it's also some description that is in fact satisfied by Clark Kent (Superman), *i.e.*, if $\hat{o} \in m(k,l,\Gamma)$, then since the second conjunct in (3.5) is true iff for *every* member o of $m(k,l,\Gamma)$, $B_l\neg F[x:o]$ is true , in particular,

$$B_l\neg F[x:\hat{o}]$$

must be *true*, that is,

Of Clark Kent (Superman), Lois believes that he *can not* fly.

So it seems that there would be contradictory *de re* belief of Lois. However, the problem shouldn't arise if we realize that, in this version, a model that correctly describes the situation must not have the two modes of presentation having any member in common. For, when Lois utters the two sentences, there is an implicit belief in her mind that Superman (as she understands the name) is not Clark Kent (as she understands the name):

$$B_l\langle xy.x \neq y\rangle(s,k), \tag{3.7}$$

for, otherwise, she wouldn't have uttered those two sentences; see the last item of Theorem 3.1.2.

Now for (3.7) to be true at the real world Γ, the model for this version is required to meet the condition that the two sets, $m(s,l,\Gamma)$ and $m(k,l,\Gamma)$ are disjoint; that is, the modes of presentation of these two names for Lois must be such that nothing satisfies both (see the remark in Section 3.1.3). So the person Superman (Clark Kent) could only fit at most one of them. Therefore, having supposed that Superman $\hat{o} \in m(s,l,\Gamma)$, we must have

$\hat{o} \notin m(k, l, \Gamma)$. So what the second conjunct of (3.5) requires is that for every member o of $m(k, l, \Gamma)$, $B_l \neg F[x : o]$ is true at Γ, which does not in any way contradict with the first conjunct. We have just one *de re* belief report, (3.6).

Again, the situation can be altered so that the person Superman (Clark Kent) does not actually fit the mode of presentation of 'Superman' for Lois, and he may or may not fit the mode of presentation of 'Clark Kent' for her. If he fits the latter, then the *de re* belief report would be: *Of* Clark Kent (Superman), Lois believes that he can not fly. If he fits neither modes of presentation, then no *de re* belief could be concluded. In any case, a model can always constructed to model the situation without any contradiction.

3.1.4 Existence

Theorem 3.1.6 $\vdash B_i\langle x.E(x)\rangle(a)$.

This theorem is to the effect that, whatever name a is, so long as a believer i has formed some understanding of it (*i.e.*, he associates with a some certain descriptions), then he believes that a exists. The validity of the theorem follows from the monotonicity condition of the domains and that $m(a, i, \Gamma) \subseteq D(\Gamma)$ for any a, i, Γ. If i has never heard of a name a, there is no mode of presentation of a for i, in which case he can not form any (*de dicto*) belief concerning it. Note that

$$\nvdash B_i E(a).$$

We observe that **B-d** seems to give the semantics of **FMP** an unnatural flavor: if a person i does not believe a is φ at Γ, that is, if $B_i\langle x.\varphi\rangle(a)$ is false at Γ, then there would have to be some member o of $m(a, i, \Gamma)$ such that i does not believe o is φ. An improvement of the semantics to overcome this disadvantage is addressed in the last chapter, the Conclusion.

3.2 Formalization of Some Well-Known Puzzles

3.2.1 Kripke's Puzzle About Belief

We deal with Kripke's puzzle involving Pierre and London in terms of the distinction of *de re* and *de dicto* beliefs. We hold that from the sentences that Pierre uttered, only *de dicto* belief can be reported, whereas the question asked by Kripke, which give rise to the puzzle, is an issue of reporting *de re* belief. The two have no necessary connection unless extra information is provided. Without the extra information, the question cannot be answered. To be more specific, from Pierre's saying 'Londres est jolie', we infer, 'Pierre believes that Londres, as he understands the name, is pretty':

$$B_p\langle x.P\rangle(Londres); \tag{3.8}$$

this is true at Γ iff for every member o of $m(Londres, p, \Gamma)$, $B_pP[x : o]$ is true, that is, iff for every object satisfying the mode of presentation of '*Londres*' for Pierre, Pierre believes that the object is pretty.

From Pierre's saying 'London is not pretty', we infer, 'Pierre does not believe that London, as he understands the name, is pretty':

$$\neg B_p\langle x.P\rangle(London); \tag{3.9}$$

this is true at Γ iff for some member $\hat{o}$ in $m(London, p, \Gamma)$, $B_pP[x : \hat{o}]$ is false, that is, iff for some object satisfying the mode of presentation of '*London*' for Pierre, Pierre does not believe that that object is pretty.

The question Kripke asked, 'Does Pierre, or does he not, believe that London is pretty?' is the question whether

$$B_pP(London) \tag{3.10}$$

is true or false. It would be true at Γ iff $B_pP[x : \sigma(London, \Gamma)]$ is true, that is, iff Pierre believes the city designated by *London* is pretty. The truth value of (3.10) cannot be determined from either (3.8) or (3.9), *unless* either $\sigma(London, \Gamma)$ is a member of $m(Londres, p, \Gamma)$ in which case we may conclude that of the city London, Pierre believes it is pretty , or $\sigma(London, \Gamma)$ is $\hat{o}$ in which case we may conclude that of the city London, Pierre does not believe it is pretty. We first may imagine a case as follows:

Suppose that whenever Pierre heard of the name 'Londres', he thinks of a city where a queen resides. In this case, since there is more than one city satisfying this description, $m(Londres, p, \Gamma)$ has more than one member, and contains London, Amsterdam, etc. Suppose that whenever Pierre heard of the name 'London', he thinks of the city where he is living at the time; thus $m(London, p, \Gamma)$ has a unique member, the city London. By (3.8), Pierre believes that all the cities satisfying the mode of presentation of 'Londres' for him—the city London, the city Amsterdam, etc.—are pretty, therefore, of the city London, he believes that it is pretty. On the other hand, by (3.9), Pierre does not believe that the city London is pretty. This clearly contradicts the above. What does this mean? It simply means that the case we have imagined cannot be a case: given that Pierre uttered those two sentences, it must be that there is at least one object which is in $m(London, p, \Gamma)$ but not in $m(Londres, p, \Gamma)$—otherwise he wouldn't have uttered both sentences: since 'being the residence of a queen' is how he understands the name 'Londres', he must have regarded it as a uniquely satisfied description, and he must not have thought the city he was living in is the same city as the residence of a queen, for then he would think the two are the same. So instead of the modes of presentations imagined above, the mode of presentation of 'London' for Pierre may be 'where buses run on the left side of the street'. This is satisfied by not only London, and if the $\hat{o}$ is, say, Hong Kong, then we wouldn't be able to infer that of the city London, Pierre does not believe that it is pretty.

In general, if a, b are two names, and if a person i utters the sentence

"a is φ',

we can conclude "of b, i believes that it is φ only if the reference of b uniquely satisfies the mode of presentation i associates with 'a' this is in fact the content of the second part of

Theorem 3.1.2, by which we have, in the Pierre-London example,

$$\vdash B_p\langle x.P(x)\rangle(Londres) \wedge B_p\langle x.x = London\rangle(Londres) \supset B_pP(London);$$

and similarly,

$$B_p\langle x.\neg P(x)\rangle(London) \wedge B_p\langle x.x = London\rangle(London) \supset B_p\neg P(London).$$

Thus, if we were to infer from the story contradictory beliefs— that Pierre believes that London is pretty and that London is not pretty, *i.e.*, $B_pP(London) \wedge B_p\neg P(London)$ we would need $B_p\langle x.x = London\rangle(Londres) \wedge B_p\langle x.x = London\rangle(London)$ to be true; but it is not true, since otherwise $B_p\langle xy.x = y\rangle(Londres, London)$ would be true, *i.e.*, Pierre believes that what he calls 'Londres' and 'London' are the same city.

If, however, one insists that according to the translation principle, on hearing Pierre saying "Londres est jolie", we ought to translate 'Londres' into 'London', and report that Pierre believes that London is pretty, then we would need to report more than that: we must also give some explanations about the circumstance, to the effect that Pierre, according to what he thinks of what 'Londres' means, believes that *that* city is pretty. It is simply because the French word 'Londres' is normally translated into 'London' that we report the situation by saying "Pierre believes that London is pretty". In the absence of any explanations, however, such a sentence would be misleading. After all, why should 'Londres' be translated into 'London'? It is because the French, based on the English spelling of 'London', have made the word 'Londres' mean whatever the English speaking people mean by 'London'; in other words, 'Londres' is meant to *denote the same object* (have the *same reference*) as 'London' does. So in translating names from one language to another, we have actually added a *de re* factor to the names. No wonder confusions arise when the distinction between *de re* and *de dicto* is taken into account.

3.2.2 Cicero vs. Tully

This example was discussed by Kripke in [25]. Suppose Jones assents to both "Cicero was bald" and "Tully was not bald", we would be justified to infer that *of* the man Cicero, Jones believes that he was bald, *provided* that the real person Cicero meets the mode of presentation Jones associates with the name 'Cicero'; alternatively, we would be justified to infer that *of* he man Cicero, Jones believes that he was not bald, *provided* that the man Cicero meets the mode of presentation Jones associates with the name 'Tully'. For example, suppose Jones has read a poem by Pope in which the name 'Tully' was mentioned; Pope was, of course, mentioning the man Cicero who's also called 'Tully', but Jones did not know the connection between these two names. Suppose whenever Jones hears the name 'Tully', he associates it with the person mentioned in the poem. Now suppose for some reason Jones believes this person was not bald, so he says: "Tully was not bald." In this case, since the man Cicero was the unique person whom Pope was referring to by the name 'Tully' in the poem, the man Cicero uniquely satisfies the mode of presentation Jones associates with 'Tully', therefore, we can report: *of* the man Cicero, Jones believes that he was not bald. If, at the same time, Jones associates with the name 'Cicero' a different mode of presentation, *e.g.*, as the most famous Roman orator, then this mode of presentation is also uniquely satisfied by the man Cicero; but then according to our logic, we would be able to derive

$$B_j\langle xy.x = y\rangle(Tully, Cicero),$$

that is, Jones believes that Tully and Cicero, as he understands the names, designate the same person. In that case, then, Jones would not have asserted both 'Cicero was bald' and 'Tully was not bald'. So *given* that Jones has asserted both sentences, (a) if we were to report the second sentence as 'Jones believes that Tully was not bald', it must be that the modes of presentation of 'Cicero' and 'Tully' for Jones are two disjoint sets, *i.e.*, no object satisfies both, (b) if we were to report the second sentence in a weaker way, as 'Jones does not believe that Tully was bald', it must be that there is at least one object which satisfies the modes of presentation of 'Tully' but not that of 'Cicero'. In either case, he doesn't think the two names name the same person, hence there is no such contradiction as "of the man Cicero, Jones believes that he was bald and that he was not bald".

3.2.3 The Imagined Case of Gödel vs. Schmidt

Kripke [24] gave an imaginary story: Gödel, who has been credited as the man who proved the incompleteness of arithmetic , actually stole the result of another man named Schmidt. Since everyone thinks, by mistake, that Gödel is the one who proved the theorem, and to most of us this is what we associate with the name 'Gödel', when we mention the name Gödel, are we referring to the man Gödel or the man Schmidt? Let's suppose that someone, Jones, always has 'the man who proved the incompleteness of arithmetic' in mind whenever he hears the name 'Gödel', and has also seen numerous pictures of Gödel and knows what Gödel looks like—in particular, he saw a picture in which Gödel was walking with Einstein in Princeton campus—and heard or read some anecdotes about Gödel. Suppose Jones says:

'Gödel was a Platonist',

shall we infer that *of* Gödel, Jones believes that he was a Platonist, or *of* Schmidt, Jones believes that he was a Platonist?

The question is not so simple to answer: it depends on what mode of presentation Jones has in mind to associate with the name 'Gödel' when he utteres the sentence. The following are possible situations:

(a) If the mode of presentation of 'Gödel' for Jones is simply 'having proved the incompleteness of arithmetic': then since Schmidt was the unique person that meets this description (that is, in our logic, the unique member of the set representing the mode of presentation), and Jones believes that whoever satisfies this description was a Platonist, we conclude that *of* Schmidt, Jones believes that he was a Platonist.

(b) If the mode of presentation of 'Gödel' for Jones is 'the man who is shown in the picture walking with Einstein': then since Gödel was the person uniquely satisfying this description, we conclude that *of* Gödel, Jones believes that he was a Platonist.

(c) If the mode of presentation of 'Gödel' for Jones is 'the man who proved the incompleteness of arithmetic *and* who is shown in the picture walking with Einstein'—this is most natural, since Jones thinks one and the same man is involved. Then since in fact there is no one satisfying this mode of presentation (the first part is satisfied by only Schmidt, the second by only Gödel), the set representing the mode of presentation is $\emptyset$, just as in the case when the mode of presentation is a contradictory description, so according to our logic, Jones would have believed anything about 'Gödel', and, moreover, since there is nothing satisfying the mode of presentation (*i.e.*, nothing is a member of the set representing it), there is nothing of which we can say "*Of* so-and-so, Jones believes he was a Platonist".

This may sound not convincing: many will argue that while Jones thinks of 'Gödel' as designating the man who both proved the incompleteness of arithmetic and was shown walking with Einstein in the picture, he doen't believe *everything* about such a 'man': he may believe something about 'Gödel' but not believe some other things about 'Gödel'. Well, in that case people (either those who argue or Jones) are confused with regard to exactly what mode of presentation Jones has in mind when uttering the sentence. While Jones believes that the same person did both things, he may be thinking of one (proving the Theorem, say) at one time and the other at another time, and no doubt there are also times when he's thinking of both when using the name 'Gödel', but then, since nothing satisfies both, even though he was talking about 'Gödel', there is in reality nothing Jones has referred to, thus we cannot attribute a *de re* belief of Jones about any object's being a Platonist based upon the sentence he uttered.

3.3 Some Unsolved Cases

In the above, we have shown that our formal system based on the notion of modes of presentation can be applied to some well-known examples in philosophy so that puzzles disappear. However, some problems remain.

We have been able to formalize sentences of the form

'i believes that a is φ',

but there are types of sentences for which such formalization does not seem to work, as in the following examples:

'Many are unaware that Cicero is Tully' (Kripke: [25]),

'Everyone believes that Madonna is musical' (Schiffer: [53]).

The problem with these sentences is that there doesn't seem to be a single sense associated with the names in question: the 'many' people have different understanding of the names 'Cicero' and 'Tully', though they are all unaware that the two names name the same person; and despite everyone's believing Madonna is musical, each person has formed the belief in his own way, depending how he understands the name 'Madonna'. In There seems to be no way to formulate such sentences in our system. The root of the problem lies in the philosophical theory on which **FMP** is based: the theory of sense and reference of Frege.

Conclusion

The logic system **FMP** shows, in a formal way, that the notion of modes of presentation enables us to deal with a considerable part of the issues about names and belief in quite a natural way (replacing modes of presentation as sets by modes of presentation as formulas would make it more natural). At the same time, since there remain several cases unsolvable by this approach, any theory of names and belief based on this notion is defective, which will certainly affect its implementation (when would a robot named Hal "think" that *everyone*—or *every* robot—believes that he is not human?). We, like many, hope that some day a more comprehensive theory will be developed which overcomes such difficulties; a formalization and implementation of such a theory would be of great interest, and Robots would be able to serve as detectives solving cases like that of Dr. Jekyll and Mr. Hyde. At present, we propose the following.

1. An immediate improvement of **FMP** will naturally be returning to our original plan, that is, instead of formalizing modes of presentation as sets, formalize them directly as formulas. The formula representing the mode of presentation of a name for a person is the description which that person associates with the name in question (cf. Section 2.1). The resulting logic will be more natural and powerful. Indeed, as a precursor of the current formalization, we did sketch such a semantics: in place of **B-d**, we would have:

   $M, \Gamma, \mu \models B_i\langle x.\varphi\rangle(a)$ iff

   $M, \Gamma, \mu \models (\forall x)(m(a,i)(x) \supset B_i\varphi(x))$.

In view of some defect mentioned above about the semantics of **FMP**, however, the following seems to be a more natural alternative:

$M, \Gamma, \mu \models B_i\langle x.\varphi\rangle(a)$ iff

$M, \Gamma, \mu \models B_i(\forall x)(m(a,i)(x) \supset \varphi(x))$;

that is, i believes that a (as he understands the name) is φ, just in case he believes that anything satisfying the description which he associates with the name a is φ.

2. The axiom system is relatively intuitive, but not well-suited for automation. One direction for future studies is to formalize, say, a tableau system, which will be a step toward automation.

Bibliography

[1] Asher, N.: "Belief in Discourse Representation Theory", *Journal of Philosophical Logic*, 15, 1986, 127-89

[2] Burge, T.: 'Sinning Against Frege', *Philosophical Review*, 88, 1979, 398-432

[3] Crimmins, M.: 'Talk about Belief', Cambridge, Mass.: MIT, 1992

[4] Donnellan, K.: 'Reference and Definite Descriptions', *The Philosophical Review*, 75, 1966, 281-304

[5] Donnellan, K.: 'Proper Names and Identifying Descriptions', *Semantics of Natural Language*, ed. Davidson, D. & Harman, G., Dordrecht: D. Reidel, 1972, 356-79

[6] Evans, G.: 'The Causal Theory of Names', *Aristotelian Society: Supplementary Volume* 47, 1973, 187-208

[7] Fagin, R., Halpern, J., Moses, Y. & Vardi, M.: *Reasoning about Knowledge*, Cambridge, Mass.: MIT, 1995

[8] Fitting, M.: *First Order Logic and Automated Theorem Proving*, New York: Springer-Verlag, 1990

[9] Fitting, M.: "Basic Modal Logic", *Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol. I, Oxford: Oxford University Press, 1993

[10] Fitting, M. & Mendelsohn, R.: *First Order Modal Logic*, Boston: Kluwer, 1998

[11] Forbes, G.: 'Indexicals and Intensionality: A Fregean Perspective', *The Philosophical Review*, 96, No. 1, 1987, 3-31

[12] Forbes, G.: "The Indispensability of *Sinn*", *The Philosophical Review*, Vol. 99, No.4, 1990, 535-63

[13] Frege, G. (1879): "*Begriffsschrift*", *Translations from the Philosophical Writings of Gottlob Frege*, ed. Geach, P. & Black, Max, Oxford: B. Blackwell, 1970

[14] Frege, G. : "On Sense and Reference", *Zeitschrift für Philosophie und Philosophische Kritik*, 100, 25-50, 1892

[15] Frege, G.: "Thoughts", *Collected Papers On Mathematics, Logic, and Philosophy*, ed. McGuinness, B., Oxford: B. Blackwell, 1984

[16] Frege, G.: *G. Frege: Philosophical and Mathematical Correspondence*, ed. Gabriel, G. et al., Chicago: University of Chicago Press, 1980

[17] Garson, J.: "Quantification in Modal Logic", *Handbook of Philosophical Logic*(ed. Gabbay, D. & Guenthner, F.), Vol. II, Boston: Kluwer, 1884, 249-307

[18] Hintikka, J.: *Knowledge and Belief, an Introduction to the Logic of the Two Notions*, Ithaca, New York: Cornell University Press, 1962

[19] Hughes, G. & Cresswell, M.: *A New Introduction to Modal Logic*, London: Routledg, 1996

[20] Kaplan, D.: 'Quantifying In', *Words and Objections: Essays on the Work of W. V. Quine*, ed. Davidson, D. & Harman, G., Dordrecht: D. Reidel, 1969

[21] Kaplan, D.: 'How to Russell a Frege-Church', *Journal of Philosophy*, 72, 1975, 716-29

[22] Kleene, S. C.: *Introduction to Metamathematics*, North-Holland, 1952

[23] Kripke, S.: "Identity and Necessity", *Identity and Individuation*, ed. M. K. Munitz, New York: New York University Press, 1971

[24] Kripke, S.: *Naming and Necessity*, Cambridge, Mass.: Harvard University Press, 1972

[25] Kripke, S.: "A Puzzle about Belief", *Meaning and Use*, ed. Margalit, Dordrecht: D. Reidel, 1979, 239-83

[26] Lambert, K.: 'Russell's Version of the Theory of Definite Descriptions', *Philosophical Studies*, 65, 1992,

[27] Linsky, L. (ed.): *Reference and Modality*, Oxford: Oxford University Press, 1971

[28] Linsky, L. (ed.): *Names and Descriptions*, Chicago: University of Chicago Press,1977

[29] Linskly, L.: *Referring*, London: Routledge & Kegan Paul, 1966

[30] Ludlow, P. & Neal, S.: 'Indefinite Descriptions: In Defense of Russell', *Linguistics and Philosophy*, 14, 1991

[31] Marcus, R.: 'Rationality and Believing the Impossible', *Journal of Philosophy*, 80, 1983

[32] Marcus, R.: *Modalities*, Oxford: Oxford University Press, 1993

[33] Mendelsohn, R.: "Frege's *Begriffsschrift* Theory of Identity", *Journal of the History of Philosophy*, 1982

[34] Neale, S.: *Descriptions*, Cambridge, Mass.: MIT Press, 1990

[35] Perry, J.: 'The Essential Indexical', *Noûs*, 13, 13-21

[36] Plantinga, A.: *The Nature of Necessity*, Oxford: Oxford University Press, 1974

[37] Putnam, H.: 'Meaning and Reference', *Journal of Philosophy*, Vol. 70, 1973

[38] Quine, W. V. O.: *From a Logical Point of View*, Cambridge, Mass.: Hardvard, 1961

[39] Quine, W. V. O.: *The Ways of Paradox*, New York: Random House, 1966

[40] Richard, M.: *Propositional Attitudes*, Cambridge: Cambridge University Press, 1990

[41] Russell, B.: *Principles of Mathematics*, Cambridge: Cambridge University Press, 1903

[42] Russell, B.: "On Denoting", *Mind,* 14, 1905, 479-93

[43] Russell, B.: "Knowledge by Acquaintance and Knowledge by Description", *The Problems of Philosophy,* Oxford: Oxford University Press, 1912

[44] Russell, B.: "The Philosophy of Logical Atomism", *The Monist,* 28 & 29, 1918

[45] Russell, B.: *Introduction to Mathematical Philosophy,* London: Allen & Unwin, 1919

[46] Russell, B.: *An Inquiry into Meaning and Truth,* London: Allen & Unwin, 1940

[47] Russell, B.: *Human Knowledge, Its Scope and Limits*, London: G. Allen and Unwin, 1948

[48] Russell, B.: 'Mr. Strawson on Referring', *Mind,* 66, 1957, 385-9

[49] Salmon, N.: *Frege's Puzzle,* Cambridge, Mass.: MIT Press, 1986

[50] Salmon, N.: "Reference and Information Content: Names and Descriptions", *Handbook of Philosophical Logic,* IV, Dordrecht: D. Reidel, 1990, 409-61

[51] Salmon, N. & Soames S. (ed.): *Propositions and Attitudes*, Oxford: Oxford University Press, 1988

[52] Schiffer, S.: "The Mode-of-Presentation Problem", *Propositional Attitudes—the role of content in logic, language and mind,* ed. C. Anderson & J. Owen, Stanford, 1990

[53] Schiffer, S.: "Belief Ascription", *Journal of Philosophy,* 89:10, 1992

[54] Schiffer, S.: "A Paradox of Meaning", *Noûs,* 28:3, 1994

[55] Schiffer, S.: "Descriptions, Indexicals, and Belief Reports: Some Dilemmas", *Mind,* 104, 1995

[56] Schwartz, S. (ed.): *Naming, Necessity and Natural Kinds*, Ithaca, New York: Cornell University Press, 1977

[57] Smullyan, A.: 'Modality and Description', *Journal of Symbolic Logic*, Vol. 12, No. 1, 1948, 31-37

[58] Smullyan, R.: *First-Order Logic*, New York: Springer-Verlag, 1968

[59] Soames, S.: 'Lost Innocence', *Linguistics and Philosophy*, 8, 1985, 59-71

[60] Taschek, W.: "Would a Fregean be Puzzled by Pierre?" *Mind,* 97.385, 1988

[61] Taschek, W.: "Frege's Puzzle, Sense, and Information Content", *Mind,* 101.404, 1992

[62] Taschek, W.: "Opacity in Context: On Ascribing Beliefs", paper presented at the Eastern Division APA meetings, Atlanta, Georgia, December 1996