Towards a Theory of Semantic Competence

Rasmus K. Rendsvig

November 2011

Supervisors
Prof. Dr. Stig Andur Pedersen
Dr. Anders J. Hede Madsen

"Humanity called them Demons without understanding what it had named"

– Neil Gaiman, *The Sandman: Season of Mist*

Contents

Abstract

In the present thesis, a theory of semantic competence is modeled using tools from epistemic logic. The resulting formal model is used to analyze a problem from the philosophy of language, namely Frege's Dilemma.

There are two aims of the thesis: to construct a formal theory of semantic competence, and to show that the formal theory can be used as an useful analytical tool in uncovering the informational structure behind problems from the philosophy of language.

The first aim is achieved by, first, deciding for which theory of meaning a theory of semantic competence is wanted. Due to its simplicity, Millianism is chosen. Then various non-formal theories of semantic competence are evaluated with respect to finding one which allows for an objective, inter-subjective comparison of competence levels. It is argued that the conceptual theory of (Marconi, 1997) is the best choice: the theory has a clearly defined structure making modeling possible, and is based on empirical studies from cognitive neuropsychology. Following these initial choices, the modeling framework and its philosophical interpretation is presented. The framework used is epistemic logic, and both the propositional and quantified versions are introduced. As a more expressible logical language is required, many-sorted quantified epistemic logic is presented, and a novel, general completeness result is shown for many-sorted extensions of quantified modal logic. Having thus set the stage of achieving the first aim, a slightly simplified version of the theory of (Marconi, 1997) is modeled. A suitable model-class is defined and a meaning function is added to capture Millian meaning. Based on the shown completeness result, a sound and complete axiom system is presented, and a logic representing the formal theory is thereby found. The model is then validated. It is shown that both the essential ontological properties as well as the competence types from Marconi's theory are present. It is further shown that the formal counterparts of the competence types from Marconi's theory adhere to the principles dictated by empirical studies. Thereby, the first aim is achieved.

To accomplish the second aim, proof of concept is shown. This is done by analyzing an objection to the correctness of the Millian theory of meaning, namely Frege's Dilemma (Frege, 1892). The formal theory is used to analyze both disjuncts of the dilemma, while focusing on the epistemic situation of the agent, i.e. the agent's level of semantic competence. The formal theory of semantic competence allows for multiple notions of semantic competence, each resulting in a unique rendering of the dilemma. Based on these analyses, it it is concluded that once the underlying informational structure of the discussed situations is revealed, neither disjunct proves to be a problem for the Millian theory of meaning. Hereby, the second aim

is accomplished.

However, I raise an intuitive objection to one of the analyses. It is argued that the objection introduces an un-accounted for parameter, namely contexts. In order to show that this objection is not fatal for the proposed analysis, a chapter is devoted to the construction of a contextual theory of semantic competence. The notion of contexts is incorporated into the models for semantic competence, and the possibilities for finding a complete axiomatic system is discussed, but no completeness result is shown. Therefore, a formal theory, i.e. a logic, for contextual semantic competence is not presented. However, the model-theoretic machinery is used to re-analyze the problematic case. It is shown that when the situation is modeled in a contextual model, the epistemic analysis of the disjunct again showed the Puzzle about Identity is unproblematic for the Millian view.

Overall, the constructed formal theory of semantic competence is shown to elucidate informational aspects of the problems posed to the philosophy of language by Frege's Dilemma. In particular, once the informational structure of the problems is clear, it is shown that each argument is far from being as decisive against Millianism as has been the mainstream view in 20th century philosophy of language.

Acknowledgements

I have been writing on this thesis for far too long — and I cannot thank my family, my friends and my wonderful Cathlin enough for their support: for warm meals I didn't have to cook, and stimulating conversations about something else.

Further, I would like to thank my supervisors Anders J. Hede Madsen from Mathematics, Roskilde University, and Stig Andur Pedersen from Philosophy and Science Studies, Roskilde University, for their support throughout.

Key parts of the thesis have been presented at various occasions, and I would like to thank the Vincent F. Hendricks for proof reading the paper accepted at the ESSLLI 2011 Student Session[1], and the anonymous reviewers for their criticisms and comments. Further, I would like to thank the participants at the ESSLLI Student Session, the IT & Cognition Guest Lectures, the Epistemic Logics for $n \geq 2$ Workshop at the University of Copenhagen and the students and faculty at Munich Center for Mathematical Philosophy for constructive criticism and enlightening discussions.

Over the next hundred and some pages, there is bound to be unconventional spelling and sub-optimal formulations. However, I am very grateful for the help from the following three people, without whom there would undoubtedly have been many more: Emil Peter Thrane Hertz helped me proof-read and revise the first three chapters, but the correction of many silly mistakes corrected throughout is ~~do~~ due to him, Stig Andur Pedersen for his thorough reading of formally dense early drafts, always spotting the missing ' ' ', and finally, Pelle Guldborg Hansen for reading through the complete manuscript in a version which in retrospect seem not much more than a draft.

<div align="right">

Rasmus K. Rendsvig
Copenhagen, November 2011

</div>

---

[1]Available in the online proceedings:
`http://www.stanford.edu/~danlass/esslli2011stus/ESSLLI2011StuS-Proceedings.pdf`

1 Introduction

Human understanding of natural language is a peculiar thing. In so many instances of everyday life, sentences are heard, read and understood. Upon understanding a phrase, information is received, knowledge is gained, and qualified decisions are made regarding which bread to buy, where to look for the keys or how to prove a complex theorem.

Even though understanding is typically taken for granted in social interaction, the notion itself is not well-understood and under debate.[1] In the present thesis, a formal theory of understanding, or, to avoid everyday connotations, of *semantic competence*, will be constructed. The guiding idea is that multiple levels of understanding are possible, and depending on how well basic words are understood, varying degrees of information is received when a statement is announced.

To give an example, imagine that we are at a social gathering, but we have run out of social lubricant. Knowing that I still have more wine in my car, you offer to fetch this, and I reply

"Yes, but the car is locked. My brother has the keys."

This plain, English statement encodes three important elements relevant to our joint desire that you get the wine. First, the request is granted. Second, an obstacle is identified, and third, a plan is proposed: collect the keys from my brother, unlock the car, and then bring the wine. Supposing these are all the relevant facts we need to consider and that my brother is present, could I now reasonably expect to sit and wait for you to return with the wine? On the one hand, yes: the statement encodes all relevant information for the goal to be obtained. On the other hand, this depends on whether or not you understood the statement. In particular, it depends on whether you understood the statement *well enough.* You may have understood the second part well enough for you to get the keys, and then the wine. Or, it may be the case that you understood the second part in a purely conceptual way: you may have understood it well enough to paraphrase it upon request, tell me that the renewed plan should then be to ask my brother for the keys or crack some joke about my family tree, but not well enough to act. How can this be? Well, it may be because you do not understand the singular term 'my brother' in anything but a conceptual way. More specifically, you may not know who the term refers to – and if you do not know the meaning of the subject in the sentence, how should you be able to follow up on the new plan?

---

[1]See e.g. (Wikforss, 2009)

This is an example where lack of understanding results in an agent less informed than it would have been given a stronger understanding of the terms of language. Often, understanding a sentence is paraphrased as 'grasping it's meaning', and the example may be seen as illustrating different strengths of 'grasping the meaning' of the singular term 'my brother'. Hence, to some degree, understanding hinges on *meaning*, which again is not an uncontroversial notion.

In fact, in 20th century philosophy, natural language meaning has been one of the main focus areas. One of the earliest contributions to the now vast literature on the philosophy of language was Gottlöb Frege's classic 1892 paper *Über Sinn und Bedeutung*[2]. The paper spawned a large scale discussion regarding the meaning of natural language terms, many issues of which are still unresolved. So far, no consensus has been reached on solutions to problems posed in Frege's paper, regarding the meaning of singular terms. This is in spite of over a hundred years active discussion among philosophers, working with a wide range of theories of meaning. Among the possible candidates are some based on *Millianism,* the theory of meaning originally criticized by Frege. These are called *direct-reference* theories. Other theories are based on Frege's original solution, and evoke an additional layer of meaning apart from mere reference, often denoted *sense* or *descriptionist* theories of meaning, see (Devitt and Sterelny, 1999; Lycan, 2000). After Bertrand Russell's detailed exposition of the first descriptionist theory of meaning in the early 20th century, such theories where widely regarded as providing an analysis ample of solving Frege's puzzles. However, this was changed by the 1972 lectures of Saul Kripke, later published as *Naming and Necessity* (Kripke, 1980). This lecture series caused a broad revival of direct-reference theories, see e.g. (Stalnaker, 1999). The ensuing debate is still unresolved, and the various arguments from both sides makes professor William G. Lycan conclude in an entry on the semantics for proper names that

> [What we are left with] is a trilemma, because it has further seemed that we are stuck with one of these three possibilities: either the names are Millian, or they abbreviate descriptions outright, or in some looser way such as Searle's, they have some substantive 'sense' or content. But none of these views [are] acceptable. (Lycan, 2006, p. 272)

The topic of the semantics of singular terms is vase, and one route to determining the correct semantics could be to expound the various arguments for and against the various meaning theories, and try to find a theory compatible with all objections. However, in the present thesis, a fundamentally different approach will be taken. The debate hinted at above has focused on finding a proper *theory of meaning* for natural language terms. Hence, the debate has focused on language itself and how it is used. In contrast, the present thesis will focus on the *language users* and their *knowledge about* the language they use. This is done by modeling agents that possess a language, and who possess information about how this language relates to the world. The focus will be on agents and their *semantic competence*. The hope

---

[2]Translated as 'On Sense and Reference', (Frege, 1980).

is that such an epistemic approach will be useful as an *analytical tool*, which can be used to shed light on the *informational structure* of certain problems from the philosophy of language. The shift in focus will be one from *theories of meaning* to one of *epistemic of meaning*: instead of focusing on what difference a change in theory of meaning makes to a problem, the theory of meaning will be kept constant while the agents' *knowledge is altered*. This shift in focus from analyzing theories of meaning to analyzing *agents' information about such theories of meaning* provides a novel, and as will be argued, fruitful approach to classic problems in the philosophy of language.

Audience

The present work is a master thesis in philosophy and mathematics with a focus on philosophical problems analyzed using mathematical methods. In particular, it is a *modeling* project, modeling cognitive aspects of language use utilizing tools from *quantified modal logic* useful for constructing qualitative models. As such, both the topics worked on, as well as the audience, divide.

One topic is philosophical. It is argued that focusing on semantic competence adds an interesting and useful new dimensions to old discussions from the philosophy of language. Whether these arguments are convincing depends on the reader's philosophical inclinations. The second topic is mathematical. In order to construct a formal model of semantic competence, a new and very expressive family of modal logics will be constructed. This family of logics is well-behaved, which a meta-theoretic analysis of them will show. The proof of the main theorem will most likely be hard to understand for most readers without a mathematical/logical background. The same can be said with regard to many of the proofs of propositions and much of the more rigorous argumentation used. The reader not interested in this is encouraged to skip it, at least on a first reading, and instead focus on the conclusions. The content of most propositions shown are of direct philosophical interest as they relay in a direct manner various features of the model.

Three distinct audiences are envisioned: one which focuses almost exclusively on the philosophical aspects, but do not care much about formalism. Such should read the thesis for its philosophical contribution, but might have a hard time understanding all arguments. A second audience is that trained in formal methods with an interest in philosophy of language or epistemology. This is the main audience, and for these readers, the thesis should introduce both all relevant formalism as well as its philosophical interpretation. The final audience are those with a formal training, but only with an interest in formal logic. They may find the philosophical arguments uninteresting, but should find all relevant definitions, lemmas and argumentation presented in order to be convinced by the proofs for the properties stated.

Aims

As mentioned above, the topics of the present work divide, but they work together towards two joint aims:

1. To construct a formal theory of semantic competence, and

2. To show that using the formal theory as an analytical tool can be useful in uncovering the informational structure behind problems from the philosophy of language.

These aims are not in opposition to any well-known position, but the epistemic logical approach is not typical in the philosophy of language.

The reason for choosing an epistemic logical approach to the topic of philosophy of language is two-fold. First, epistemic logic is an active and fruitful research area. A lot of research has been done recently in epistemic logic, rational interaction and the dynamics of information exchange, see e.g. (Baltag and Smets, 2008; van Ditmarsch et al., 2008; Grossi et al., 2009, 2010). Here, the qualitative modeling style of epistemic logic has been successful in uncovering various informational aspects of belief revision, learning, awareness, trust, questions and answers and more. Second, many of these topics are close to those present in the philosophy of language. As the possession of a language is a requirement for communication, it is natural to ask what knowledge rational agents may have of their language, and to which degree this affects their understanding of their language. So far, this question that has not yet been addressed in the literature on logic and rational interaction.[3] That the approach is not mainstream in the philosophy of language, it is conjectured, is a matter of the field traditionally focusing on theories of meaning in conjunction with the relatively young age of the paradigm of epistemic logic.

Approach   In order to obtain the first aim, a formal theory of semantic competence is constructed by utilizing epistemic logical tools to model a conceptual theory of lexical semantic competence based on empirical studies from cognitive neuropsychology. The formal theory gives rise to various notions and degrees of semantic competence, differing in the amount of information possessed by the agent in question. In fact, in the theory it is easy to clearly distinguish between strengths of semantic competence relevant for the example regarding 'my brother' above. Such distinctions allows for an analysis of what information an agent can extract from a statement, as this depends on the agent's level of understanding.

The formal theory is identified with *a logic*. Yet, the fundamental approach will be model-theoretic, and a completeness proof is therefore required in order to specify the appropriate logic. The model-theoretic approach focuses on formula's *truth* and *validity* relative to a model or set of models. In contrast, the core notion in relation

---

[3]As mentioned in the acknowledgements, parts of this thesis has been presented at various occasions, with senior researchers from the field present. Yet, the author was not referred to works taking a similar approach.

to logics is whether a formula is *provable.* A completeness result for a class of models and a logic establishes the fact that anything *valid in all the models* can be *proven in the logic.* Hence, where the model-theoretic structure validates the appropriate formulas relative to the conceptual theory and the completeness result is proven, a logic capable of proving all these validities is obtained.

The second aim is obtained by showing *proof of concept.* This is done by applying the theory to a classical problem from the philosophy of language, namely *Frege's Dilemma.* Various versions of the problem is formalized using the concepts of the theory, and it is shown how this gives a clear picture of the informational structure. The problems are modeled as restrictions in the set-theoretic models, and it is shown that given precise notions of semantic competence, they are compatible with the theory of meaning they were originally used to reject. The detailed analyses of the problems provides novel and intuitive solutions based on the uncovered informational structure.

Limiting the project

Validation.   As mentioned above, a formal theory/model of semantic competence will be constructed and used. This model is to a high degree based on the conceptual, non-formal model of (Marconi, 1997), which in turn is based on empirical studies. Hence, the modeling construction taken as a whole compromises three distinct elements, namely 1) empirical findings supposed to reflect reality, 2) a conceptual model elucidating the important aspects of the collected data, and 3) a formal model which ought to capture the important aspects of the conceptual theory, and the logical consequences of which should fit further collected data.

In order to properly validate the model, all three points should be addressed. However, this is beyond the scope of the present work. For review of the empirical studies and the arguments in favor of the conceptual model, the interested reader is referred to (Marconi, 1997). The reader interested in learning about the assumptions and methods of cognitive neuropsychology is referred to (Coltheart, 2001).

Regarding the third point, this is done half-way, insofar as the model is constructed to fit the properties of the conceptual model, but is not compared to further empirical studies. Some properties which would allow for such a comparison is derived once the formal theory has been constructed, and theoretic possibilities for falsifying the model are discussed in section 5.5.2. A comparison with further studies is an interesting venue for further research, but is far beyond the scope of the present work.

Connections with philosophy.   In this thesis, work is done on the epistemics of understanding the meaning of words across possible worlds. Hence, all major elements of analytical philosophy are touched. Unfortunately, it is beyond the scope of the present work to relate it to all the various debates in epistemology, metaphysics, the philosophy of mind and the philosophy of language that could seem interesting. As a consequence, a lot of assumptions are made without comment. For example is

neither semantic nor epistemic internalism/externalism discussed, though it should be clear from the text that externalism is adopted. Neither is epistemic logic related to epistemology in general, and the usage of epistemic alternatives and multiple contexts not related to metaphysics. The basic assumptions of epistemic logic are discussed, but not in relation to the general literature on the philosophy of mind. The epistemic logical framework used as a modeling tool double as an epistemological framework, and as such stands on it's own. It is assumed that the paradigm of epistemic logic is sound. These limitations are due to the complexity of these many topics and the vast literature on these. An attempt to situate the present work to these discussion would be to open Pandora's box, and would delay the constructive approach used to the point where nothing would be accomplished.

The core concept in the thesis is semantic competence. This is often seen as a part of linguistic competence, which includes both semantic and syntactic competence. Of these two, syntactic competence has received a lot of attention both in the philosophy of language and in linguistics, see e.g. (Devitt, 2006). The notion of syntactic competence will not be discussed in any but a trivial sense: agents are assumed and modeled as syntactically competent with respect to single words, which is taken to mean that they always know when two name tokens are of the same type.

Taking this into consideration, the focus of the thesis is rather narrow. The pragmatic, epistemological framework of indistinguishability as embedded in quantified S5 epistemic logic (see chapter 3) is taken as a proper modeling tool for certain aspects of knowledge, and this is related to a possibly naive interpretation of human concepts, cognition and semantic competence. No problems regarding direct cognitive connection with objects are discussed; it is assumed that the agents are capable of identifying objects, and in some philosophically unspecified way represent these cognitively. It is assumed that the agents are able to categorize the world they inhabit. There will be no discussion of skepticism, only an information-based notion of knowledge with the pragmatic goal of action. Language, words and meanings are taken to be something which agents can have knowledge about, but only a very limited range of language is modeled. Meaning is taken to be truth-conditional, external, and is for each word assumed to be given by a function. All these elements limit the scope of the project, hence making the subject feasible for the study of a master thesis, but are of course open for discussion.

Outline

The thesis is structured as follows. Meaning and semantic competence are treated in chapter 2. The Millian view of names is presented along with various objections, among which are two raised by Frege in his classic paper. Following is a section on semantic competence. In order to use such a theory to compare competence levels of agents, the theory must conceive of semantic competence in a way that allows for this. A selection of theories are presented and discussed in relation to the requirements. The conceptual theory of (Marconi, 1997), which will be modeled in chapter 5, is introduced among these. In chapter 3, the epistemological framework

and modeling tool, epistemic logic, is introduced. This chapter includes two main sections: one on propositional epistemic logic and one on quantified epistemic logic. The propositional case is less mathematically challenging, but includes the main philosophical ingredients: the possible worlds framework and the indistinguishability notion of knowledge. Once the mathematics and philosophy of the propositional case has been expounded, the system is extended to the first-order case, quantified epistemic logic (QEL). The mathematics of QEL are more complicated than the propositional case due to the introduction of constants, predicates and quantification. Each require a formal interpretation, but also a philosophical one. In this respect, notions of concepts and object identification are in focus. As the modeling of Marconi's conceptual theory requires an augmented version of QEL, the fourth chapter is devoted to a strict introduction of many-sorted modal logic. This chapter is highly technical, and includes the main mathematical result of the thesis: a theorem allowing for easy proofs of completeness for a large family of normal, many-sorted modal logics. The theorem is variant of the Canonical Model Theorem well-known from the literature on modal logic, see e.g. (Blackburn et al., 2001). One of these many-sorted modal logics will be used to model Marconi's theory, which is introduced in chapter 5.[4] The chapter outlines some simplifying assumptions regarding Marconi's theory and this simplified version is then modeled. It is shown that the formal model includes the appropriate ontologies and can express the different competence types from Marconi's theory by modeling these as different information states of the agents. The formal theory is throughout compared with Marconi's theory, and it is shown that it captures the important features. Furthermore, the formal model is discussed with respect to properties not discussed by Marconi. Notably, more levels and competence types can be identified in the formal framework, providing a more fine-grained taxonomy of semantic competence levels. In chapter 6, the model constructed is applied in order to show proof of concept. This is done by analyzing the two Fregean puzzles introduced in chapter 2. The conclusion of both puzzles depend on the agent's understanding of natural language terms. The puzzles are modeled in the constructed framework and evaluated by utilizing the strict definitions of semantic competence obtained in the previous chapter. These analyses, focusing on the epistemic state of the agents, provide natural solutions to the puzzles. In one case, an intuitive objection is raised. The core insight of the objection is that object identification often is context dependent. In order to show that once this parameter is included in the model a satisfactory analysis and solution is still provided by the epistemic approach, the model is augmented with contexts in chapter 7. This structure is discussed from a logical and a philosophical viewpoint, and an epistemic analysis of the puzzle is carried out, showing how the objection silently includes an appeal to lacking cross-identificational knowledge. In the final chapter, multiple cases from the philosophy of language where an epistemic approach can be applied is outlined, and a general conclusion is drawn.

---

[4]The chapter is accessible to readers who skipped chapter 4.

2 Meaning and Semantic Competence

To relate world, language and agency, two important notions from the philosophy of language will discussed in this chapter: *meaning* and *semantic competence*. The viewpoint on meaning discussed in the present thesis is that of *Millianism*. An introduction to Millianism and objections is the topic of the first section. Among the objections is *Frege's Dilemma*, an argument presented originally in Gottlöb Frege's *Über Sinn und Bedeutung*. Both disjuncts of the dilemma are argued by Frege and other, see e.g. (Collin and Guldmann, 2010; Lycan, 2006), to be unfeasible. Therefore, the dilemma is seen to be a definitive argument against Millianism. The dilemma and the disjuncts, denoted *Frege's Puzzle about Identity* and *The Problem of Non-Informativeness*, are introduced below and later analyzed in chapter 6 using a formal theory of semantic competence.

In order to construct such a formal theory, a conceptual starting point is required. Therefore, the second section is devoted to the introduction and discussion of various theories of semantic competence from the philosophical literature.

Due to the foundational character of the work, there will throughout be a focus on the most basic terms from naturals language, namely *singular terms*. Singular terms encompasses most notably proper names, like 'Hans Reichenbach' and 'The Morning Star', but the use of 'My brother' in the example from the introduction is also a singular term. The latter is also an instance of an *indexical* – a term changing its meaning relative to the speaker. Though many indexicals are prominent, singular terms, for simplicity such will not be considered.

Singular terms provide a natural starting point as these are among the most basic terms of natural languages. They present the least complicated terms with respect to meaning and reference, cf. (Collin and Guldmann, 2010). Further, they occur in the most fundamental subject/predicate constructions and in almost all aspects of daily communication and information transfer.

2.1 Millianism and Frege's Puzzles

To first focus the relation between world and language, this section discusses *meaning*. Demarcations can be drawn between various general viewpoints on the meaning of natural language sentences and the proper way to analyze such. Most notably, the theories of the 20th century have been focusing either on *use* or *truth*, cf. (Loar, 2006). Here, the focus will be on truth, and the tradition of *truth-conditional seman-*

tics will be followed. As the arguments against Millianism rests on this approach, a short introduction will be provided.

The core idea of truth-conditional semantics is that the meaning of the morphemes[1] of a given language, when organized into sentences, determine the meaning of the sentence in a systematic way that allows one to settle whether the sentence is true or false. A clear example of a language with adopted truth-conditional theory of meaning is a first-order language $\mathcal{L}$ with semantics defined as usual, see e.g. Hendricks and Pedersen (2011, ch. 4). Adopting a truth-conditional theory of meaning for natural languages hence means assuming that this approach can be used to analyze the sentences of natural language.

When defining the truth-conditions for formulas in first-order logic, the meaning of the non-logical signs are given purely *extensional*. That is, the full meaning of a constant is given by an object in the domain of interpretation, and the meaning of a predicate is also given by it's extension, a subset of the domain. The meaning of the logical constants are taken to be functions from formula/truth-value pairs to truth-values. The meaning of 'and', '$\wedge$', for example, would be the function

$$
\begin{aligned}
((\varphi, \top), (\psi, \top)) &\longrightarrow \top \\
((\varphi, \top), (\psi, \bot)) &\longrightarrow \bot \\
((\varphi, \bot), (\psi, \top)) &\longrightarrow \bot \\
((\varphi, \bot), (\psi, \bot)) &\longrightarrow \bot
\end{aligned}
$$

Here, φ and ψ are first-order formulas of the language $\mathcal{L}$ and $\top$ and $\bot$ denote, respectfully, truth and falsehood.

Truth-conditional theories of meaning have the intuitive strength of straightforwardly relating lexical items and sentences to the world via the "simple" notions of truth and reference. As an underlying framework for meaning theories, it hence expresses in a "simple" manner the informational content of a sentence (possibly relative to a given context).

In the example above, it may seem that a truth-conditional approach entails that meaning is expounded only in terms of extension/reference. This is not the case. While the semantics used for first-order logic works in exactly that way for the non-logical signs, other theories of meaning are compatible with the truth-conditional view. It is typically argued, though, that reference lies at the heart of meaning, and that this aspect must be considered in order to construct a feasible theory of meaning. As case in point is Devitt and Sterelny (1999), who refer to this viewpoint as *the representational thesis*. Arguments aiming to gain support for this thesis typically revolve around the use of language to convey information about the world. In order for terms to be useful in conveying information about the world, they must in some way stand in a relation to the world, and a simple hypothesis is that this relation is that of *reference*. Working in the tradition of truth-conditional semantics

---

[1] The smallest meaning carrying items of the language. For example proper names, logical connectives and certain prefixes – like 'un-'.

further forces one to accept this thesis as the non-logical terms *must* refer for truth to be determined.[2]

Truth-conditional semantics does not dictate that reference must be the *only* aspect there is to meaning. In order to solve his puzzles (see below), Frege added an additional layer, namely 'sense'. The literature on this is vast, and will not be considered as the main focus is not on theories of meaning, but instead on the agents' information.

In order to utilize any type of truth-conditional semantics, *some* theory of meaning for the morphemes must be chosen. In the following, the simplest such, *Millianism*, will be expounded.

2.1.1 Millianism

Named after John Stuart Mill, who famously presented the theory in his 1843 book, *A System of Logic*, the core claim of Millianism is that the meaning of a word is exhausted by its referent. Hence, the semantics for first-order logic discussed above may be seen as an instance of a Millian theory of meaning: the full meaning of a constant is its extension.

As a general theory of meaning, Millianism may seem obviously insufficient. As Lycan (2000) argues, for example, the logical connectives seem not to denote anything.[3] Or, if words are merely names of objects, then sentences are merely strings of names: but strings of names does not, in any obviously way, possess meaning like sentences do.[4] Hence, Millianism for all words may seem implausible, but it does seem to have a certain merit when the discussion turns to singular terms.

Taking the "Millian view" of names, as Devitt and Sterelny (1999) calls it, implies that names work merely as labels attached to objects.[5] For this reason, the theory has also been referred to as the *label theory* (Collin and Guldmann, 2010). The view states that the meaning of a given, unambiguous proper name is constituted solely by the object to which the name refers, i.e. by its referent. On this view, the meaning of the name 'Hesperus' (the Evening Star) is constituted by the planet Venus construed as an existing object, and nothing more. Hence, strictly speaking, the meaning of the name *is* the referent of the name.

One strong argument in favor of Millianism is that it is a *minimal* theory of meaning. It assumes only one aspect of meaning, namely reference, and as this aspect is required by the representational thesis mentioned above, it cannot be given up unless truth-conditional semantics are abandoned.

---

[2]Depending on ones notion of *truth*. In the present, a *correspondence theory of truth* will be assumed without further argument. For an introduction, see (Collin and Guldmann, 2010).

[3]One could argue that logical connectives refer to *functions*, as mentioned above, but this discussion is for the present irrelevant.

[4]Lycan (2000) invites us to consider the string 'Fred Martha Irving Phyllis' and remarks that even if two of those names denote abstract properties, it would still not be seen as meaningful as normal sentences are.

[5]Theories regarding how names are attached to objects are know as *theories of reference*. The possible choices for such will not be discussed.

2.1.2 Frege's Dillemma

Unfortunately, the argument above apparently in favor of Millanism can also be framed as one of the most prominent arguments against it. The specific counter-argument is often referred to as *Frege's Puzzle about Identity*, cf. (Lycan, 2000), and originates from Gottlöb Frege's classical 1892 paper *Über Sinn und Bedeutung*[6]. The problem, as Lycan presents it, is one disjunct of a dilemma Frege puts to Millianism, cf. (Collin and Guldmann, 2010).

    The dilemma revolves around the way one should understand identity statements. Assume that 'n_1' and 'n_2' are two names from a language which also includes an identity relation '=', and form the sentence '$n_1 = n_2$'. This compound expression can be ambiguous. As Collin and Guldmann puts it,

> [The] identity relation expresses a relation between objects or between signs that refer to objects. (Collin and Guldmann, 2010, p.49)

If the first disjunct is adopted, *Frege's Puzzle about Identity* results. If the second disjunct is opted for, a problem which will be denoted *The Problem of Non-Informativeness* results. These will now be presented in turn.

Frege's Puzzle about Identity   If one chooses the first disjunct in the above-mentioned dilemma while holding a Millian theory of names, one may be confronted with the following argument which implies an inconsistent theory. Consider the two true identity statements

(a) Hesperus is Hesperus

(b) Hesperus is Phosphorus

Given that the two names co-refer[7], all terms in the two identity statements have the same meaning, and hence (a) and (b) must have the same meaning from a truth-conditional point of view. Therefore, according to Millianism, (a) and (b) must be equally informative to a semantically competent speaker of English. As the first is a trivial validity of self-identity, this does obviously not carry informational content. Opposed to this, the latter seems to be a contingent, empirical fact, and hence convey information. *If* this is true, (a) and (b) *do* differ in informational content, and the Millian view should be rejected. The argument can be pinned out as follows:

(A) (a) and (b) mean the same.

(A→B) If (a) and (b) mean the same, then a semantically competent speaker would know that (a) and (b) mean the same.

---

[6]'*On Sense and Reference*', listed as (Frege, 1980).

[7]'Phosphorus' or 'the Morning Star' does in fact denote the planet Venus. So does 'Hesperus' and 'the Evening Star'.

(B→C) If a semantically competent speaker would know that (a) and (b) mean the same, then they are equally informative to the speaker.

(¬C) (a) and (b) differ in informativeness to the competent speaker.

∴ Contradiction.

The four premises are jointly inconsistent, and the typical textbook choice.[8] is to reject premise (A). This premise is a consequence of the Millian view, and the conclusion drawn is that there must be more to meaning than mere reference.

The Problem of Non-Informativeness The first disjunct of Frege's Dilemma seems to be inconsistent with the Millian view, and the Millian is therefore forced to opt for the second disjunct. This, however, is equally problematic, as will now be argued.

If the second disjunct of the dilemma is chosen, then, where n_1 and n_2 are proper names,

> the informational content of the sentence '$[n_1 = n_2]$' consists in the sentence's expressing the fact that the sign '$[n_1]$' designates the same object as the sign '$[n_2]$'. (*ibid.*, p.49-50)

This may be seen as problematic, as the identity statement no longer conveys information regarding non-linguistic reality, but rather about linguistic conventions. It is argued in (Collin and Guldmann, 2010) that this disjunct too is unfeasible, as identity statements will be vacuous. The problem arises, they argue, since being informed that n_1 refers to the same object as n_2 does not provide information about word-world relations unless it is already know which object n_2 refers to. It is concluded that

> [u]ntil we do, we are [in possession] of an item of information about two languages and not knowledge about the world. (*ibid.*, p.50)

Though not a knock-down argument, the counter-intuitive conclusion of this argument forces one, it is argued by Frege, to choose the first disjunct, which in turn drives us out of the "Millian paradise", to borrow an expression from (Devitt and Sterelny, 1999).

2.1.3 Further Objections

There are many further objections to the Millian view, even when this is restricted to only proper names. These objections are introduced here as they will become relevant in chapter 8, where they will be discussed in relation to venues for further research. The objections originate with Frege and Russell. The formulations presented here stem from various textbooks, see e.g. Lycan (2000); Collin and Guldmann (2010); Gundersen (2003).

---

[8]See, e.g., (Collin and Guldmann, 2010), (Lycan, 2000) or (Gundersen, 2003).

The Problem of Substitutivity   The Problem of Substitutivity revolves around *Leibniz' Law* or *Substitution of Equals* when used in relation to belief ascriptions. As an axiom of first-order logic, this states that for some formula with occurrences of a, denoted $\varphi(a)$, and the same proposition with occurrences of b replacing all those of a, denoted $\varphi(b)$, if $a = b$ then $\varphi(a)$ is provable if, and only if, $\varphi(b)$ is provable. That is, the two are logically equivalent, so whenever one has an occurrence of the first and the true identity statement, one can obtain the other via logically valid deduction. In section 3.2.4, this principle will be referred to as the *Principle of Substitution*:

$$(x = y) \rightarrow (\varphi(x) \leftrightarrow \varphi(y))$$

The problem posed to Millianism, then, is that this fails in intentional, '*opaque*' contexts, like that of belief. Consider the proposition 'Lois Lane believes that Clarke Kent wears glasses'. In the fictive universe of Joe Shuster og Jerry Siegel, this proposition would probably hold true. In the same universe, so would the identity statement 'Clarke Kent = Superman', and thus, by Leibniz' Law, so should 'Lois Lane believes that Superman wears glasses'. But this is a false belief ascription, and the obtained sentence is therefore false. Hence, as Leibniz' Law preserves truth-values when substituting names that mean the same with one another, but the exemplified proposition changes its truth-value, there must be more to the meaning of names than mere reference.

The Problem of Empty Names   In the example just given, names of fictive characters was deliberately used. Supposedly, the reader had no problems understanding the paragraph. Given the Millian view of names, this is quite mysterious. For how could any sentence involving names of such non-existent entities as 'Superman' be understood when part of the components in it literally have no meaning? Given that the meaning of 'Superman' is the referent of the name, and this referent does not exist, it must follow that the name *has no meaning*. Hence, it may be concluded that there must be more to the meaning of names than mere reference, for otherwise 'Superman' would be meaningless.

The Problem of Existence Statements   Closely related to the problem of empty names is that of existential statements, i.e. statements that claim the existence, or lack thereof, of the bearer of a name. Consider the statements 'Superman exists' and 'Superman does not exist'. These two statements are the negations of one another, and by the law of the excluded middle[9], one of them should hold true.

As the first statement is meaningful, which it is by 'intuition', then it must be true – for on the Millian view, only names that refer have meaning, and the statement could not be meaningful if one of the terms in it where not. But this violates reality, because Superman does not exist. However, looking at negated statement,

---

[9]The law of the excluded middle says that for all propositions p, either p is the case, or p is not the case, where the latter is understood as stating that *not-p* is the case. If this disjunction is assumed in a logic, one can conclude from the falsity of p to the truth of *not-p* and vice versa.

this, too, is meaningful. But if it is meaningful, then not only is it false, it is a logical contradiction. But this conclusion is too strong. For though its falsity is in agreement with current facts, it is still a metaphysical possibility that Superman *could have* existed in some other possible world. But then this world would be violating the laws of logic. Such paradoxes cannot be tolerated, and the only possible conclusion is that there must be more to meaning of names than mere reference.

In what follows As mentioned in the introduction, Frege's Dilemma will be used as case studies for an epistemic analysis, and solved in order to provide a proof of concept. Alas, constructing and introducing a framework capable of analyzing all the above objections is beyond the scope of the present work. Once the formal theory for semantic competence has been introduced and applied, possible extensions capable of modeling these objections are addressed as venues for further research (see chapter 8). For the analysis of the disjuncts in Frege's Dilemma, a more simple framework is required, as the modeling required does not have to include terms for non-existing objects or beliefs in the agents' language.

The four objections caused a widespread dismissal of Millianism as a proper theory of meaning for names throughout the 20th century. As an alternative, Frege suggested a dual aspect theory of meaning, where the meaning of names consisted in both their reference and *sense* (an objective 'mode of presentation'). Elaborating on this theory, Russell introduced his description theory, according to which the meaning of a name was a *definite description.* Various critiques raised later caused Searle to introduce a variant of the theory, where a single definite description was replaced by a cluster of such.[10] These dual aspect theories were in turn challenged by Kripke (1980), arguing that names function as *rigid designators*, meaning that they denote the same individual in all metaphysically possible worlds. This hypothesis is inconsistent with the mentioned dual aspect views, and caused a wide rejection of these, as well as a revival of variants of Millian theories of meaning, often denoted *direct reference* theories.

The discussion on the semantics of singular terms after Kripke is nicely summarized by the quote of William G. Lycan in the introduction, stating that a trilemma seems to be in effect. Solutions to the trilemma have been constructed, see Lycan (2006) for references.

In the following, a different strategy will be adopted. In the hope that a bloated ontology of more complex meaning theories can be avoided, the focus is changed from theories of meaning to theories of semantic competence. To show the viability of this approach, Frege's Dilemma will be analyzed in light of the *information* the agents possess about their language and the way this language is related to the world they inhabit.

---

[10]See e.g. (Lycan, 2000) for an introduction to these theories and the various objections.

2.2 Semantic Competence

Semantic competence is in general not a well-defined term, and its usage is far from standardized. Often, the notion is used without explicating its meaning, and even in articles seemingly focusing explicitly on the topic, the notion is left undiscussed (see e.g. Callaway's *Semantic Competence and Truth-Conditional Semantics* (1988)). The notion will here be used as an information comparison device between agents regarding their knowledge of the language they use. As such, it is used as an objective measure, which allows for objective, inter-subjective comparison of said competence. This means that the sought for a theory must give clear criteria for whether an agent is semantically competent or not, where the verdict of competence is independent of those how pass it and the situation in which it is passed.

Such a use is in contrast to the view of semantic competence used in (Rast, 2006), which depends on both subjective status and social context. Rast argues that

> successful identification of the semantic referent is not generally a requirement by virtue of linguistic competence. (p. 37)

Rast draws this conclusion as he can produce situations of successful communication where the parties involved are unable to identify the referents of proper names. The requirements put to speakers in Rast's setting is highly context dependent. In one example, Alice working in customs is competent when stating "The RMS St. Helena has a capacity of 128 passengers", though she is unable to identify the relevant ship. "In this case," Rast writes, "she might not be required to be able to identify the ship in question, simply because it is not part of her job." (p. 39). Besides being context dependent, Rast's use is highly subjective: an example is cited where a history professor is unable to identify George Washington on a one dollar bill and is indirectly deemed incompetent with respect to the name *due to his expertise*, whereas other speakers without said expertise in the same context may be deemed competent with respect to the name without the ability to perform the required identification.

In contrast to this view, where semantic competence is a matter of peer judgment relative to subject and context, the notion of semantic competence used in the present is intended as an objective one. This notion is to be used to construct a theory of semantic competence which provides clear-cut characteristics, that allows for inter-subjective comparison of competence levels. Such a theory is sought in order to investigate how the competence level of language users affect the arguments of the preceding section.

It should be noted that it is assumed that a satisfactory notion of semantic competence *simpliciter* cannot be found. Rather, it is assumed that agents will be semantically competent *with respect to* some part of language – be it a language, a sentence or a set of sentences or lexical items. The theories regarded here focuses on sentence and lexical semantic competence, with the previous being more coarse-grained. This will be regarded first, where-after the more detailed cases are examined.

2.2.1 Truth-Theoretic Semantic Competence

A rather rough theory of semantic competence may be derived from the Fregean/Tractatus-Wittgensteinian[11] view of meaning and understanding. This rests on the core thesis of truth-conditional semantics, namely that a sentence's meaning can be equated with its truth conditions. Understanding a sentence is then equated with knowing its truth conditions, or, as Wittgenstein puts it in *Tractatus Logico-Philosophicus*[12]

> 4.024   To understand a sentence in use means to know what is the case
>   if it is true.

On this view, being semantically competent with respect to a sentence S of some language $\mathcal{L}$ with truth conditions T is equivalent to knowing that

$$S \text{ is true (in } \mathcal{L}) \text{ iff}^{13} \ T$$

To illustrate using a classic example, then semantic competence with respect to the German sentence 'Schnee ist Weiß' amounts to knowing that

$$\text{'Schnee ist Weiß' is true (in German) iff snow is white}$$

Such bi-conditional statements are well-known from the two prominent figures in 20th century philosophy, namely Tarski and Davidson. The former finds that such T-sentences (or T schema's) can fully describe the meaning of formulas from formal languages (Tarski, 1972), as is common practice today, whereas the latter found that a theory of meaning for a natural language would be a finitely axiomatized theory yielding as theorems T-sentences, one for each sentence in the natural language, where the truth-conditions are purely extensional (Davidson, 1984) – a thesis know as *Davidson's Program.*

A stronger version of the view may be formulated by incorporating knowledge across possible worlds, as is done by Lycan when he writes:

> It is often said that a person P knows the meaning of a sentence S if P knows S's truth-conditions, in the sense that given any possible world (or possible situation), P knows whether S is true in that world (or situation). (Lycan, 1994, p. 203)

This latter formulation sets stricter standards for language users in order for them to be semantically competent with respect to a given sentence, as the former only requires knowledge in one possible world, namely the actual, whereas the latter

---

[11]Exactly who has supported this view, and to what degree, is beyond the scope of this thesis. Two articles criticizing similar views are (Lycan, 1994) and (Soames, 1989), but neither are explicit about *who* holds these views.

[12]The quote is adopted from Wiggins (1999, p. 5).

[13]'iff' is used throughout as an abbreviation of 'if, and only if'

requires knowledge across all. Both views have been criticized in (Lycan, 1994, ch. 9) and (Soames, 1989).

Both formulations has certain merits for the present purpose. In particular, they allow for a inter-subjective comparison of competence with respect to sentences based on a strict measure. In the weaker formulation, it may, based on agent knowledge of T-schema's, be decided whether the two agents are equally competent with respect to a given sentence. The stronger version allows for the same comparison regarding an agent's knowledge across the set of possible worlds. Combining the weaker notion with the possible worlds framework would allow for a hierarchy based on the cardinality of the set of possible worlds in which the agents are competent in the weaker sense.

Yet, one can argue that both views ultimately only allow for a bivalent judgment regarding sentence competence per possible world. In fact, any theory which has sentences as the smallest element with which agents can be competent is flawed in the sense they allow for a less fine-grained division of levels of semantic competence than theories that focus on word (lexical) competence. To see this, assume that two agents are both *in*competent with respect to a simple subject/predicate sentence. In this case, a sentence-based theory will characterize them both as being equally incompetent with respect to the sentence. A lexical theory will be able to judge *why*: one agent may be incompetent with respect to the subject, whereas the other may be incompetent with respect to the predicate or both. Hence, the sentence-based approach would in this case provide one category, whereas a lexical-based would provide three.

Further, exactly due to its focus on sentences, such theories can be criticized regarding learnability. As argued by Davidson (1984), any language and theory of meaning must consist of a finite set of morphemes and a finite set of syntactic and semantic rules in order for its infinitely many sentences to be learnable by a finite agent. Such requirements are not in thread with either of the above theories of semantic competence. These theories in fact precludes the possibility of semantic competence with respect to natural languages: natural languages can produce an infinity of sentences and hence no finite agent can know a T-sentence for each. This insight further strengthens the idea of a theory of semantic competence which takes a lexical focus.

In fact, an alternative source (Wittgenstein, 1922) for the above Wittgenstein quote suggests a partial lexical focus in his viewpoints as well:

> 4.024   To understand a proposition means to know what is the case, if
> it is true.
> (One can therefore understand it without knowing whether it is
> true or not.)
> One understands it if one understands its constituent parts.

Here, the last line indicates a relation between sentence and lexical understanding, though which relation depends on how one reads the latter implication. As the present thesis is not meant as an excursion in interpreting Wittgenstein, this will

not be dwelt on. Instead, the topic will be changed to theories focusing explicitly on *lexical* semantic competence.

2.2.2 The Translation-Ability view

To find a theory of semantic competence which allows for an objective, inter-subjective comparison of language users, lexical theories seems to provide a degree of detail not present in sentence based theories. One theory of semantic competence which takes a lexical focus is presented in (Devitt and Sterelny, 1999). Here, the notion is embedded in a general framework of linguistic competence, which consists of syntactic competence and lexical competence.

Notably, the viewpoint presented by Devitt and Sterelny does not consist in propositional knowledge, but is ability based:

> Understanding a language no more involves having propositional knowledge of a semantic sort about the language, or representing its rules, than being able to ride a bicycle involves having propositional knowledge of a mechanical sort about riding, or of representing the mechanics of riding. (Devitt and Sterelny, 1999, p. 187)

On their view, competence in using a language is an ability to utilize the constructs of the language to express thoughts which have the appropriate meaning in relation to the language used. Further, it is an ability to relate constructs of the language to thoughts which have such appropriate meanings. From this it is concluded that linguistic competence requires a certain amount of conceptual competence. In particular, agents must be able to entertain thoughts with the appropriate meanings. Devitt and Sterelny accept the *Language of Thought Hypothesis*, and this conceptual competence is spelled out in terms of competence with respect to *Mentalese*[14]. Specifically,

> competence in the language *requires* the competence to think Mentalese sentences with meanings expressible in the language. (p. 188)

Based on the assumption of competence with Mentalese, syntactical and lexical competence is regarded by Devitt and Sterelny as two translational abilities. For syntactical competence, this requires the ability to translate between sentences from the given language and sentences of Mentalese such that the appropriate sentences have "structure similar enough to count as translations." (p. 188). It is further required that the agent is syntactically competent with respect to Mentalese, which consists in the ability to construct Mentalese sentences based on the Mentalese lexicon. Syntactical competence will not be discussed further.

---

[14]In one sentence, the Language of Thought Hypothesis states that all thinking is done in a physically realized language in the brain of the subject. The language of thought is often referred to as Mentalese. The correctness of the hypothesis will not be discussed, but for an introduction, see (Aydede, 2010).

Devitt and Sterelny does not discuss semantic, lexical competence much. For such competence, they hold that the agent must be able to translate back and forth between lexical items and their corresponding Mentalese counterparts. A more detailed description of what such a translational ability consists in is not discussed, and regarding it's acquisition it is only noted that this is an ability an agent has "in virtue of being part of the causal network for [the word]." (p. 189).

They further state that in order to be genuinely competent with the language, any agent must also be competent with respect to the appropriate Mentalese lexical items. What is required for this competence depends on which of at least two types of words is in question.

For basic words, the meaning of which are given by direct reference theories, competence with their Mentalese counterpart consists in "having thoughts that are appropriately linked, either directly or indirectly, to [the referent]". (p. 189). Here, having directly linked thoughts to the referent will be perceptual confrontations with the referent, whereas indirect links are by reference-borrowing and the like.

Regarding "the least basic words", which have their referents determined by a description theory of reference fixing, competence consists in associating appropriate words to one another and drawing appropriate inferences – for example associating the Mentalese 'bachelor' to the Mentalese 'male', and inferring in Mentalese 'x is male' from 'x is a bachelor'.

The exposition from (Devitt and Sterelny, 1999) of this last kind of competence does leave a lot to be wished for. Important key notions are given only a superficial treatment, as for example competence with Mentalese lexical items and how this ability is acquired. For the present purposes, this is a weakness of the theory. In particular, it is not easy to see how one may judge an agent competent or incompetent with respect to a certain Mentalese word (or mental concept, if the language-of-thought hypothesis is discarded). Experts regarding certain immunological terms will supposedly be deemed competent with respect to their Mentalese counterparts, but as agents' thoughts may also be indirectly linked to the referent through for example a causal network and reference borrowing, any child using 'cytotoxic lymphocyte' while intending to use it as his father from biomedicine will have appropriately linked thoughts, and hence be deemed competent. Hence, this notion of competence is quite vague, which in turn makes inter-subjective comparison difficult.

However, the theory does evade the problem of learnability posed above against the truth-theoretic account of semantic competence. It is sufficient for an agent to learn syntactic rules along with the meanings of morphemes in order to be linguistically competent with respect to the given language.

Finally, it's worth noting that the theory allows for two different aspects of semantic competence, though this is not explicitly commented on by Devitt and Sterelny. These two aspects are on the one hand that of acquaintance with objects as required for competence with basic words, and on the other hand issues regarding the relations among words, allowing for analytical inference. This distinction is not unique to Devitt and Sterelny, though, as will be seen in the following section.
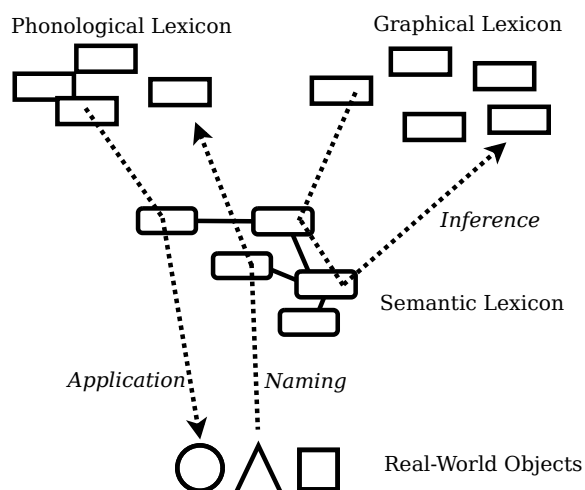
Figure 2.2.1: Marconi's theory of the structure of lexical competence. Dotted lines represent instantiations of competence types. The illustration is an altered version of (Marconi, 1997, fig. 3.2, p. 72).

2.2.3 Marconi and Lexical Competence

A more specific theory of semantic, lexical competence can be found in (Marconi, 1997). Here, Diego Marconi constructs a conceptual theory of the structure of such competence, based on studies in cognitive neuropsychology[15]. As will be argued, this theory is the one regarded best suited for the present purpose, and will subsequently be the theory modeled. In (Marconi, 1997, Ch. 3), the theory is regarded as a theory of the Structure of Lexical Competence. The proposed structure will be referred to as the SLC.

The elements of the theory consist of three relations defined over three ontologies: *real-world objects*[16], *word lexica* and the *semantic lexicon.* Each of the three relations correspond to a competence type. These are *inferential competence* and two types of *referential competence*, being *naming* and *application*, see Figure 2.2.1.

Word Lexica   The word lexica are sets of words given by some perceptual mode. In particular, Marconi mentions the *phonological lexicon* and the *graphic lexicon.* The phonological lexicon is a set of auditive perceived words, where-as the graphic lexicon is a set of visually perceived words. Though not mentioned by Marconi, further lexica could be imagined, for example a tactile lexicon had by people who can read Braille.

Each word lexicon consists of the words familiar to the subject, supposedly when

---

[15]For the review of these studies, arguments for the structure and references to relevant literature, the reader is referred to (Marconi, 1997, ch. 3).

[16]Marconi does not discuss the metaphysical nature of real-world objects, and neither will such a discussion be presented here. It is merely assumed that a world of objects does exist and that these exhibit common-sense properties such as temporal duration.

the words are known by the given mode of perception. One must assume the latter as it obviously can be the case that a subject may have a large phonological lexicon but a small to non-existent graphic lexicon – as exemplified by children who have not yet learned to read.

Semantic Lexicon   The semantic lexicon is "a collection of mental representations distinct from the representations of both phonological and graphic word forms." (Marconi, 1997, endnote p. 71). Though Marconi is far from explicit about the exact nature of these mental representations, it will be assumed that the semantic lexicon consists of the non-linguistic concepts possessed by the subject. It will be assumed that these concepts coincide with those modeled in epistemic logic, see section 3.2.3.[17]

Interesting discussions from the philosophy of mind present themselves regarding the exact nature and existence of such concepts, but these will not be dealt with. It will merely be noted that humans classify worldly objects and do have mental representations of this classification.[18] Where a given class consists of multiple objects, this will be denoted a *concept*. Where it only encompasses a single object, it will be denoted an *individual concept*. It is assumed that the concepts Marconi has in mind coincides with this use.

Inferential Competence   Inferential competence is a knowledge-based ability to draw semantic inferences regarding connections between words. This ability is founded in knowledge of true sentences involving the relevant words.

Such a true sentence with respect to the words 'apple' and 'fruit' would be 'all apples are fruits'. Marconi illustrates the concept by an example that invokes a bookish zoologist: though the zoologist may never have been in contact with a certain species of butterflies, he may nevertheless be inferentially competent with respect to its name as he would know many facts of the species.

In broad terms, inferential competence is

> the ability to manage a network of connections among words [through the semantic lexicon], underlying such performances as semantic inference, paraphrase, definition, retrieval of a word from its definition, finding a synonym, and so forth. (*ibid.*, p.59)

Inferential competence performances involve two of the three ontologies, namely the word lexica (or one such) and the semantic lexicon. Marconi states that one can

> [describe] the kind of performances in which inferential competence is typically displayed ... as following word-word routes *through* the semantic lexicon. (p. 71)

---

[17]This is a non-trivial assumption, but it is beyond the scope of this thesis to dive into the intricacies of mental representations and concepts beyond what the discussion to come of the interpretation of epistemic logic.

[18]Again, this is a non-trivial assumption, and the previous note can be restated.

This is illustrated in Figure 2.2.1. The mental gymnastics exercised in applying inferential competence is thus a process occurring in the semantic lexicon: it is in the semantic lexicon that the relation between the mental representations exist, and no such relation exist in the word lexicon. Notably, *no* such relation exist within a word lexicon.

Based on Marconi's description, see pp. 69-71, an inferential performance seems to require three steps: 1) having knowledge of a lexical item which connects this to a (individual) concept in the semantic lexicon, 2) having a mental representation of a relationship (concept) between this concept and another concept, and 3) having knowledge connecting the mental representation of the relationship and the latter concept with each their lexical item, resulting in knowledge of a connection between three lexical items, for example 'Socrates is human'.

Referential Competence   Referential competence, in turn, is "the ability to map lexical items onto the world" (p. 60), a knowledge-based ability involving all three ontologies consisting of two distinct subsystems.

The first of these subsystems is that of *naming.* Naming is the act of retrieving a lexical item from the appropriate word lexicon when presented with an object. It is assumed that this act, if successful, is carried out by 1) having knowledge of a lexical item connecting this to a concept in the semantic lexicon and 2) knowledge connecting this concept to a real-world object. It should be noted that the degree of successfulness of the act depends in both steps on the concept invoked in the semantic lexicon. In particular, the act may be less successful if this concept does not individuate an object. In relation to epistemic logic, this is discussed in section 3.2.3.

The second subsystem is that of *application.* Application is the act of identifying an object when presented with an object. Again, this is a two-stage process, which requires 1) knowledge connecting a real-world object to a concept in the semantic lexicon, and 2) knowledge connecting this concept with a lexical item from the appropriate word lexicon. The same note as above may be applied here. That the notions of ambiguity of concept from epistemic logic are invoked is warranted by the interpretation of this as object identification – or, as Marconi puts it, object recognition. The issue will be discussed in section 3.2.3. Both processes of referential competence are illustrated in Figure 2.2.1.

Empirical Reasons for Multiple Lexica and Competence Types   Marconi's SLC may seem overly complex. It may be questioned, for example, why one should distinguish between word and semantic type lexicas, or why referential competence is composed of two separate competence types, instead of one bi-directional. The reason for distinguishing between two types of lexica, as well as for individuating the three types of competence, is based on empirical studies from cognitive neuropsychology. These studies reviews of subject with various brain-injuries, and indicate that these modules of human cognition are separate. In (Marconi, 1997), literature from cognitive neuropsychology is reviewed in support of this distinctness.

In the terminology of (Coltheart, 2001), impairments of the three competence types show *double dissociation. Dissociation* between two impairments are said to exist when patients showing one may not show the second. When other patients show dissociation in the reverse direction, a *double* dissociation is said to exist. It is argued in (Marconi, 1997) that inferential competence is double dissociated from referential competence, and that the two types of referential competence are double dissociated.

The hypothesis that the word lexica and the semantic lexicon are distinct are supported by cases where patients are able to recognize various objects, but are unable to name them (they cannot access the word lexicon from the semantic lexicon). In the opposite direction, cases are reported where patients are able to reason about objects and their relations when shown the objects, yet unable to do the same when prompted by their names (i.e., the patients cannot access the semantic lexicon from the word lexicon). The latter cases provides support for the hypothesis that reasoning is done with elements from the semantic lexicon, rather than with items from any of the word lexica.

Now, Marconi's study is from 1997, so newer publications from cognitive neuropsychology may show that the proposed structure of lexical competence to be wrong. It is beyond the scope of this thesis to review the literature of cognitive neuropsychology since the publication of (Marconi, 1997). The consistency of the architecture proposed with newer empirical research can therefore not be evaluated.

2.2.4 Theory Comparison and Choice

The theories from the above sections will now be compared, and it will be argued that Marconi's SLC is best suited as the basis for a theory of semantic competence that allows for objective, inter-subjective comparison of language users.

The theories of Devitt and Sterelny and Marconi differ vastly from the truth-theoretic theory of section 2.2.1 in that they both include the lexical aspects. Yet, in structure, the former seem to resemble one another: both suggest a duality in the cognitive structure, one part consisting of natural language words, the other of mental entities. Words are related to mental entities, and these entities are in turn related to the world. The concrete mechanisms of these relations vary. In particular, where the exposition of Devitt and Sterelny is philosophically heavily theory-laden, Marconi's SLC is based on empirical data with few philosophical assumptions.

Based on the exposition above, it can be seen that both have a dual notion of semantic competence. Marconi directly analyses competence in terms of referential and inferential competence, whereas Devitt and Sterelny focus on competence with different word types, depending on the way they conceive these words have been given meaning. The two aspects they utilize are ultimately similar to referential and inferential competence: basic competence with basic words is similar to referential competence, where-as competence with "least basic words" is similar to inferential competence. Devitt and Sterelny require only one type per word, whereas Marconi always includes both types.

The general structure proposed by both has clear merits as the outline for a formal theory of semantic competence, but due to the more detailed presentation of Marconi's viewpoint, this will be the conceptual model that will be focused on in the present thesis, for the following reasons.

First, the cognitive structure of language using agents, including the relationships between the various competence types, is made clear in Marconi's SLC, and this with empirical backing. The shown relationship between the various modules makes the modeling easier, in that less guesswork will be involved. Moreover, that the structure is based on empirical data allows for a comparison of the model with further data in subsequent works.

Secondly, that the competence types and their roles are more clearly defined makes the notions of the SLC easier to apply to agents than the theory of Devitt and Sterelny. In particular, it is easier to determine whether or not an agent is competent in a certain respect, which in turn allows for easier inter-subjective comparison. Specifically, it is easy to test whether an agent is referentially competent of either type, as it merely requires to 'ask' the agent to identify the word/object in question.[19] It is further easy to decide whether an agent is inferentially competent with respect to a word by similar tests.

It should be noted though, that the notion of inferential competence has theoretical problems in relation to empirical testing. In order to decide whether a subject is inferentially competent with some word, the subject is required to perform semantic inferences, paraphrase, define and find a synonym for the word presumingly until the tester is happy that the subject knows the inferential meaning of the word. But *the* inferential meaning of a word is notoriously hard or impossible to define: this is very close to the problem of finding a principled basis for cluster description theories of meanings for proper names, as discussed by Devitt and Sterelny (1999, p. 51). This problem will be circumvented when modeling inferential competence. The modeling will be done for a language containing only proper names and identity, and in that case, precise definitions of inferential competence *with respect to a name* can be given. As the competence types can be clearly applied for each word in the modeled framework, it will hence further be possible to make precise inter-subjective comparisons for competence levels.

Finally, the precision of the structure presented by Marconi makes the SLC interesting to model as the modeling to a higher degree can be validated. In particular, the structure must be captured so it is possible to express the various competence types. Further, the modeling must respect the constraints regarding the relationship between competence types.

In chapter 5, a modeling of a slightly simplified version of Marconi's structure of lexical competence, will be undertaken. This will build on the logical framework of quantified epistemic logic, introduced in the following chapter.

---

[19]In a formal setting by checking the knowledge of the agent, and in an empirical testing, by actually asking the subject – perhaps sufficiently many times to eliminate lucky guesses.

3 Epistemic Logic

As mentioned in the introduction, the modeling tool used in the present thesis doubles as an epistemological framework: it both serves as a mathematical modeling tool, and as an epistemological theory. In this chapter, both these elements of epistemic logic will be presented. In the first section, the system S5 for propositional epistemic logic is introduced. The general conception of knowledge of the S5 framework is introduced alongside the definitions of the logical apparatus: syntax, semantics and a Hilbert-style axiom system. In the second section, syntax, semantics and axiom system for first-order epistemic logic with constant domain semantics and non-rigid terms is presented. The main discussion in this section relates to the interpretation of predicates and constants, and agents' view of these. The purpose of the chapter is hence to introduce both the logical framework and its epistemological underpinnings, which forms the basis for the later modeling.

To accommodate a less technically orientated audience, the writing style, definitions and notation used in the present chapter is less rigid then the one used in the following chapter that presents the completeness proof mentioned in the introduction. The use of parentheses to indicate scope is kept to a minimum to aid readability for the less formally orientated readers, and important, but technical, definitions are not presented. If the reader finds that anything in the present lacks clarity, all relevant definitions etc. can be found in the next chapter.

3.1 Propositional Epistemic Logic

There are many different systems of epistemic logic, each with its own axiomatic base and (possibly) corresponding set of suitable models. The most widespread system is S5. The system is a classic modal system often used to model possibility and necessity. It has also been given an epistemic interpretation, which (Fagin et al., 1995) was influential in making a mainstream modeling tool. The system can further be given a clear philosophical interpretation in both the propositional and the first-order case, and thereby serves as a proper epistemological theory. This is the interpretation of epistemic logic presented here, which to a high degree stems from the works of Jaakko Hintikka.[1]

---

[1]It is interesting to note that in the first book length piece on epistemic logic, namely Hintikka's 1962 *Knowledge and Belief: An Introduction to the Logic of the Two Notions*, Hintikka rejects one of the axioms of S5, namely 5, the axiom of *negative introspection* (see below) on intuitive grounds.

The fundamental idea behind the S5 notion of knowledge is a *negative* one based on *possibility elimination* resulting in a *partitioning* of such possibilities. Stated simplistically, it is knowledge gained through the process of elimination. The intuitive idea is that agents try to decide what is the actual world they are in – that is, what is true about the world they are in – by means of eliminating possible worlds using their available information. Eliminating worlds from being the actual one results in a partitioning of all possibilities into those that are consistent with current information, and those that are not. An agent is said to know a proposition if the proposition is true in all the possibilities consistent with the information. This may be seen as a negative view on knowledge: an agent knows proposition p, only in the case where the agent can eliminate all *non-p* possibilities. Hence, to gain knowledge of the true, the agent should eliminate the false.

Apart from this view of knowledge, no further definition of the term 'knowledge' is identified by the epistemic logical framework, cf. (Hendricks, 2006).

As the epistemological interpretation of the system is most clearly reflected by the associated model theory, this will be introduced first.

3.1.1 Epistemic Models and their Philosophical Interpretation[2]

For both model theory and syntax for a modal logic[3], two elements are fundamental. First, a countable set of *proposition symbols*:

$$Prop = \{p_1, p_2, ...\}$$

Each proposition symbol is understood as a shorthand for a non-compound proposition. For example, the proposition symbol p_1 could be a shorthand for the proposition 'The Earth is round', and p_2 stand for 'The Moon is made of cheese'. Hence, each proposition symbol is a shorthand for a true or false statements about some factual relationship.

The second fundamental element is a finite set of *agents*:

$$I = \{1, 2, ..., n\}$$

An agent is used as a generic term for any entity capable of action. This may be a human being, an animal, a computer program or the timer on an oven. In the following, action is understood broadly, and includes such static, cognitive actions as having knowledge or considering something possible. Agents are typically rational, idealized, cognitive agents.

An epistemic model based on such a set of propositions and a set of agents consists of three elements. The first is a countable set of *possible worlds*,

$$W = \{w_1, w_2, ...\}$$

---

[2]The following exposition of the semantics of propositional epistemic logic is largely based on (Fagin et al., 1995) and various textbooks on modal logic in general, see e.g. (Blackburn et al., 2001; Chellas, 1980).

[3]A strict definition of a *logic* is presented in the next chapter. Stated loosely, a logic is a set of formulas closed under a set of inference rules.

Of these, one is called *the actual world.* The interpretation is that the actual world reflects the state of affairs as they truly are. The remaining worlds are considered *alternatives* to the actual world. This means that they are logically possible other ways in which the actual world could have been. In the epistemic interpretation, these are called *epistemic alternatives.* Each possible world can be taken to be a complete description of all factual matters in some metaphysical variant of the real world, or they may be taken to be descriptions of some idealized situation which one wishes to analyze, including only aspects relevant for the modeling process. In the present work, the latter, pragmatic approach to possible worlds is taken, assuming no metaphysical commitment as to their existence. The terms 'state', 'scenario', 'situation' or 'situation description' may hence be more suitable than 'possible world', and these will all be used interchangeably throughout.

The second element of an epistemic model is an *interpretation*, which is a function

$$\mathcal{I} : W \longrightarrow \mathcal{P}\left(Prop\right)$$

that assigns to each possible world an element of the power set of the set of propositions. The intuitive idea is that the interpretation assigns to each world the set of propositions which are *true* at that world.

The third and final element of a epistemic model is a set of *indistinguishability relations*[4]:

$$\left(\sim_i\right)_{i \in I} = \left\{\sim_1, \sim_2, ..., \sim_n\right\}$$

There is exactly one indistinguishability relation for each agent i, and this is the relation $\sim_i$. Each is a binary relation between worlds, that is for each $i \in I$,

$$\sim_i \subseteq W \times W$$

Each is assumed to be an *equivalence relation*, which means that it partitions the set of possible worlds into partition cells, often denoted *information cells* in the epistemic interpretation. Within each cell, all worlds are connected to each other, but to no worlds outside the cell. Two partitions are illustrated in Figure 3.1.1.

Each information cell consists of a set of worlds, and the interpretation is, that the agent in question cannot tell these worlds apart, i.e. they are indistinguishable to the agent. Hence, if two worlds w and w' are related by agent i's indistinguishability relation, that is, if $(w, w') \in \sim_i$, or using a more pleasant notation, if $w \sim_i w'$, these worlds are said to be *indistinguishable to i*, or we say that w' is an *epistemic alternative to w for i*. That is, based on the information available to i at w, i cannot tell whether the actual situation is w or w'.

To give an example: assume agents i and j are away on a weekend trip, and i knows of j that a) he always wears exactly one pair of same colored socks, and that b) he brought one red pair and one black pair. This is all the information available

---

[4]The indistinguishability relation is often, in modal logic, referred to as the *accessibility relation.* The former term is chosen here due to the specific epistemic application and role of the relation in mind.
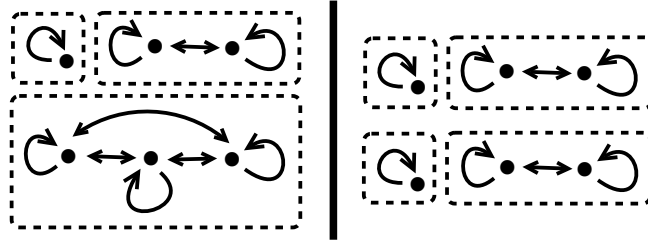
Figure 3.1.1: Two partitions of a set with 6 elements. On the left side, the partition consists of three cells, marked with dotted lines. In each cell, all elements are connected to each other, but to no elements outside the cell. On the right side, the partition consists of four cells.
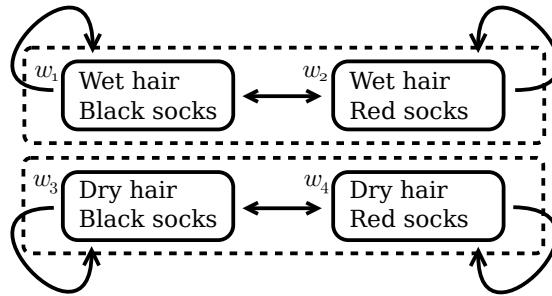
Figure 3.1.2: The indistinguishability relation is drawn with arrows. There are the two information cells: $\{w_1, w_2\}$ and $\{w_3, w_4\}$. In w_1, the agent knows 'Wet hair', but does not know 'Black socks' or 'Red socks'.

to i regarding j's garments when j exits the bathroom dressed on Saturday morning. In this case, i will know certain things; what color shirt, pants and shoes j is wearing, for this is plainly visible; whether j's hair is wet or not, as i will be able to tell these cases from each other – but i will not know whether j is wearing the red or the black socks, for this fact will be concealed (by shoes and pants). Thus, i is incapable of telling the scenario where j is wearing the red socks from the scenario where j is wearing the black socks. i cannot distinguish between the two based on the available information, and hence they are related by i's indistinguishability relation. This is illustrated in Figure 3.1.2.

This indistinguishability analysis of the relation between worlds is made possible by the semantics of S5 logics, namely that the relation must be an equivalence relation. An equivalence relation on a set divides the whole set into disjoint (not overlapping) subsets – the information cells of Figure 3.1.1.[5]

Another example of an equivalence relation is the relation 'same height as'. If applied to some population, the population will be partitioned into subsets; the set of all those of height 100, the set of all those of height 110, etc. If two from the population are the same height, they will be in the same subset, but if they are not, then they will not. In the same way the set of scenarios is partitioned into the

---

[5]I.e., an equivalence relation induces a partition on the given set.

aforementioned information cells, relative to each agent.

Thus, in the weekend trip example, there may be one set of worlds where j has wet hair, all of which are related by i's indistinguishability relation with respect to the fact of the color of j's socks, and another set of worlds where j has dry hair, as illustrated in Figure 3.1.2. If one world belongs to the first set, and another world belongs to the second set, then i will be able to tell them apart by virtue of the fact that i can tell whether j has wet or dry hair. They are thus not connected by i's indistinguishability relation. On the other hand, if two worlds belong to the same cell of the partition, the agent will not be able to tell them apart due to lack of information.

In the formal models, worlds are related by an indistinguishability relation for some agent as a matter of courtesy from the modeler's side; for instance, if the epistemic workings of an agent unable to tell day from night or wet hair from dry is to be modeled, the appropriate worlds may simply be stipulated as being related.

A formal definition of epistemic models is presented in section section 3.1.4 below.

3.1.2 Propositional Syntax

Propositional epistemic logics are propositional logics augmented with one or more modal operators for knowledge, or *knowledge operators*. For each i from the set of agents $I = \{1, 2, ..., n\}$, K_i will be a knowledge operator. Given these and a set of propositions $Prop = \{p_1, p_2, ...\}$, the set of well-formed formulas can be defined as follows: where $p_k \in Prop$, K_i is a knowledge operator and φ is a well-formed formula, the following will be well-formed formulas:

$$p_k \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \to \psi \mid \varphi \leftrightarrow \psi \mid K_i\varphi.$$

A knowledge operator K_i is read 'i knows that...'. It can hence be expressed that, for instance, 'i knows that φ and ψ' by $K_i(\varphi \wedge \psi)$. A *dual* operator of K_i may also be defined as a shorthand for $\neg K_i \neg$, which will be written P_i – i.e. for some formula φ, $P_i\varphi$ is short for $\neg K_i \neg\varphi$.[6] This dual operator is read 'it's possible, for all i knows, that...' for reasons that will become apparent shortly.[7] In the following, both K_is and P_is will be referred to as *knowledge operators*.

In case of the example above, letting p_1 read 'j is wearing red socks' and p_2 read 'j is wearing black socks', it would follow that agent i knows that one or the other of the two propositions is true. Hence, the model should make the formula $K_i(p_1 \vee p_2)$ true. Agent i knows neither that it is the one or the other, so the model should also make $\neg K_i p_1$ and $\neg K_i p_2$ true. By the definition of the dual operator, from the first of these cases it follows that $P_i \neg p_1$, i.e. it is possible, for all that i knows, that the friend is *not* wearing red socks.

---

[6]Of course, for meta-theoretical reasons, the binary connectives could be defined from either one of the binary connectives and the negation. This is the approach taken when proving completeness in chapter 4.

[7]Alternatively, the operator can be read 'φ is consistent with agent i's information', cf. (van Benthem et al., 2011).

3.1.3 Axioms and the present conception of knowledge

In the present section, the axioms of S5 will be presented and justified as fair principles of knowledge for idealized, rational agents and this notion of knowledge will further be discussed with respect to its areas of application.

The normal modal logic S5 includes the axioms of propositional logic and the following four axiom schemes and two rules of inference, where φ may be any well-formed formula and K_i is any knowledge operator as specified above:

$$
\begin{aligned}
\text{K:} \quad & K_i\left(\varphi \to \psi\right) \to \left(K_i\varphi \to K_i\psi\right) \\
\text{T:} \quad & K_i\varphi \to \varphi \\
\text{4:} \quad & K_i\varphi \to K_iK_i\varphi \\
\text{5:} \quad & \neg K_i\varphi \to K_i\neg K_i\varphi
\end{aligned}
$$

The rules of inference are *Modus Ponens*: If φ and $(\varphi \to \psi)$ are theorems, then ψ is a theorem, and *Knowledge Generalization*: If φ is a theorem, then $K_i\varphi$ is a theorem.[8]

The logic S5[9] is defined as the smallest set containing all propositional tautologies and the above axioms, and which is closed under the stated inference rules. This system is sound and complete with respect to the models and semantics of section 3.1.4 below, cf. (Chellas, 1980).

Depending on the way the two dual operators are defined, a further axiom is required for soundness and completeness, namely

$$
\text{Dual:} \quad K_i\varphi \leftrightarrow \neg P_i\neg\varphi
$$

Given the definition of the P_i operator above, this axiom is not required, but is valid on all the specified models. It is often included as an axiom, and it captures an important aspect of knowledge as understood in the S5 epistemic framework, as will be discussed below.

Discussion of each of the axioms as principles of knowledge is warranted. The first axiom, K, also known as *Distribution* or the *axiom of deductive cogency* (Hendricks, 2006, p.84), is in one crucial aspect both the most fundamental, but also the most problematic, of the S5 axioms. The axiom yields the property that any agent knows the logical consequences of their knowledge. It may be argued that this is a crucial property of strong rationality. Unfortunately, the agents comes to know *all* the logical consequences of their knowledge, and are hence *logical omniscient*. As the axioms of propositional logic are included in S5, then by Knowledge Generalization (KG), Modus Ponens (MP) and K, the agents in question will know all the theorems of the propositional calculus. As this set of theorems is infinite, even just adding KG and K to the propositional calculus results in a system far too strong to model human cognition.

---

[8] This last rule of inference is also often referred to as *Necessitation*.

[9] or S5$_n$ if one wishes to include reference to the amount of agents. The subscript will be omitted where it will not cause confusion.

For this reason the knowledge operator was originally given another reading in the first book-length piece on the subject, namely Jaakko Hintikka's 1962 *Knowledge and Belief: An Introduction to the Logic of the Two Notions.* Here, Hintikka proposes reading $K_a p$ as

> It follows from what a knows that p (Hintikka, 1962, p.31)

after having discussed problems regarding human agents' knowledge of the consequences of their knowledge and consistency reinterpreted as defensibility, cf. (*ibid.*, pp.25-31). According to this epistemological story, the theorems of epistemic logic are to be understood as rationally indefensible, where

> [i]ndefensibility is fleshed out as the agent's epistemic laziness, sloppiness or perhaps cognitive incapacity to realize the consequences of what he in fact knows (Hendricks, 2006, p.89)

There are several ways of dealing with the problems of logical omniscience, cf. (Meyer, 2001), but here neither will be invoked, and it is noted that the agents are thus treated as highly idealized with respect to their epistemic abilities. The problem will be discussed in relation to model validation in chapter 5.

The interpretation of the validity of *Dual* rests on the entire viewpoint of knowledge as understood within epistemic logic. As stated earlier and as will be made explicit in the formal semantics, knowledge in the epistemic logical framework is a matter of *possibility elimination*. If an agent can eliminate all situations in which φ is the case, then the agent is said to *know* $\neg\varphi$. Further, if the agent is unable to eliminate one or more situations in which φ is the case, then, for all the agent knows, φ may be the case, or the agent *does not know* φ. Thus, if an agent knows φ, then there are no possibilities left to consider where $\neg\varphi$ is the case, and hence it is *not* possible, for all the agent knows, that $\neg\varphi$ is actually the case – which is exactly the wisdom contained in *Dual*.

The axiom T*,* often called the *axiom of truth* (Hendricks, 2006, p.85), reflects the commonly accepted requirement of knowledge as also seen in the *justified true belief*-definition of the concept, cf. (Sosa, 1992, Tripartite definition of knowledge), that no matter what, what is known must be *true.* Put differently, T encodes the requirement that knowledge is *infallible.* As before, this axiom perhaps requires too much of the agents if knowledge should be less than ideal: before some proposition p can be said to be known, *all* possibilities of error must be eliminated, leaving only situations in which p is actually the case. Thus, unless certain assumptions are made with respect to the space of possible worlds allowed in the models, the possibility of knowledge is quickly ruled out by Cartesian skepticism. Such skepticism can have the effect that no formulas $K_i\varphi$ will be true for non-trivial φ, as it may be impossible to eliminate certain worlds: in particular, worlds in which one is being severely fooled by intuitions regarding daemons and computers. For whether such states are in fact the case may be globally under-determined by *any* available evidence *ex hypothi*, cf. (Hendricks, 2006).

Of the two final axioms, *4* and *5*, only *5* is necessary to include in the axiomatic base, as *4* can be proven by means of *5* and T. *4* is included as it embodies an important aspect of rational agency, namely that of *positive introspection*. That the agents dealt with in the present are positively introspective means that if an agent knows something, then he knows that he knows it (which is again known, etc.).

This feature of the epistemic abilities of the agents involved is strikingly implausible when compared to an everyday use of the verb 'to know' – as an example, if upon receiving the answer to a question one replies 'Ah, I knew that!', then this agent has not been epistemologically rational in the sense of epistemic logic, for this answer shows that there was no reason to ask the question in the first place; if the agent knew the answer before he received it, then he also knew that he knew it, and thus would have no need to seek the information in the first place.

An argument in favor of *4* as a reasonable principle of the elimination conception of knowledge is the following. Assume that agent i is in a situation in which he can eliminate all possibilities of a certain event, say $\neg p$. Then he will know p, i.e. $K_i p$. In this case, i will also have eliminated all possibilities in which he does not know p, i.e. all situations in which $\neg K_i p$ holds. This is so as $\neg K_i p$ holds *only if* there is some non-p world that has not been eliminated. But such have been eliminated by assumption. So, i can eliminate the possibility that he does not know p, and all that is left in situations where $K_i p$ is the case. Hence, this is known by i, i.e. $K_i K_i p$ holds.

The final axiom of S5, namely *5*, is also an introspection property, but this time of *negative introspection*. That agents are assumed to have this ability is a strong assumption, letting the agents know their own *ignorance*. This requires a full overview of all the possibilities given some state of affairs, and certainty as to which can be eliminated on the basis of current information. With respect to human knowledge of models incorporating enough facts, this is indeed implausible to assume, but when regarding certain less complicated scenarios, the principle can be seen as fair.

This, for example, is true in many card game situations where players know of their own uncertainty regarding a card deal. To exemplify *5*, recall the weekend trip example above. That i knows that j has brought only two pairs of socks limits the amount of possibilities, namely to two. In either of these two situations, there will be uncertainty as to what color socks are being worn and all situations in which there is no uncertainty can be eliminated (a situation with no uncertainty with respect to the socks color would, for instance, be one in which j was wearing no pants and shoes, but such scenarios can be eliminated by appeal to current information, i.e. the information that the friend steps out of the bathroom fully dressed). As all situations in which i knows the color of the socks are eliminated, only situations in which this fact is not known are left. And that the socks color is not known in all these states in turn makes it the case that *this* fact is known, that is, it is known that it is not known, what color the socks are.

In sum, S5 can hence be seen as the logic of knowledge for (ideal) agents possessing a complete overview of the different possibilities available.

As noted, the notion of knowledge employed in standard epistemic logic is not

based on any particular definition of knowledge, such as justified, true belief. That knowledge must imply truth is assumed in the S5 system due to the axiom T, not due to an interpretation of the knowledge operators. This in turn means that knowledge does not imply that any kind of justification is present for the propositions known, and nor does standard epistemic logic include any machinery to deal with information gathering or learning. This, in turn, means that there is no focus on the methods used to acquire information in the first place. Especially with respect to direct object constructions and *wh*-knowledge[10], this can make the reading of certain formulas unintuitive, as will be commented on further below.

3.1.4 Truth Conditions

So far, semantic structures in terms of worlds and indistinguishably relations, etc., and a language for talking about these structures have been introduced. In order to assign meaning to that language, the topic now turns to *truth conditions*, which serve exactly this role.

As mentioned earlier, for a logic based on a set of propositions *Prop* and set of agents I, any models have three basic constituents; a set W of worlds, a set of indistinguishability relations $(\sim_i)_{i \in I}$ and an interpretation $\mathcal{I} : W \longrightarrow \mathcal{P}(Prop)$. A model may hence be defined as a triple

$$M = \langle W, (\sim_i)_{i \in I}, \mathcal{I} \rangle$$

Though models allow for the determination of the truth or falsehood of some formulas,[11] it will typically be interesting to evaluate formulas at a specific point in a model, for example the actual world. For this, the notion of a *pointed model* is required.

Where M is a model and $w \in W$, a pointed model is any pair (M, w). In the following, the parentheses will be omitted.

Based on the definition of a pointed model, the truth conditions for propositional epistemic logic can now be given. Strictly speaking, *truth*[12] is defined as a relation between (pointed) models and formulas. The relation is denoted '$\models$', and the expression '$M, w \models \varphi$' is read 'φ is true at world w in model M'. The truth relation is defined by the following iff clauses:

---

[10]Knowing *wh*o, *wh*at, *wh*ere, etc., someone is, for example.

[11]In the same way *frames* allow for the determination of the validity of certain formulas. For completeness results, frames, defined as pairs $F = \langle W, (\sim_i)_{i \in I} \rangle$, are the crucial part of the models as axioms characterize properties of the indistinguishability relations, cf. (Blackburn et al., 2001). This point will be returned to in the following chapter.

[12]Also often denoted *satisfaction*, cf. e.g. (Blackburn et al., 2001). Where a formula is true, it may also be said that it is *satisfied* or that it *holds*. These terms will be use interchangeably.
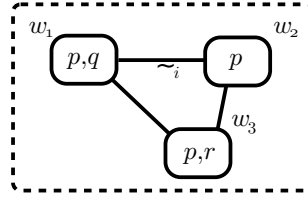
Figure 3.1.3: An example of a model with $W = \{w_1, w_2, w_3\}$. The set of formulas in $\mathcal{I}(w_k)$ is drawn in the world's box. For this model it is the case that $M, w_1 \models p \wedge q$, $M, w_2 \models \neg q$ and $M, w_3 \models p \rightarrow r$. Agent i's indistinguishability relation is illustrated by a line connecting worlds. Lines to the same world has been ignored. Hence, $M, w_1 \models K_i p$ and $M, w_1 \models K_i(p \vee q)$. Further, $M, w_1 \models \neg K_i q$, but $M, w_1 \models P_i q$ (and much more).

$$
\begin{array}{lll}
M, w \models p & \text{iff} & p \in \mathcal{I}(w) \\
M, w \models \varphi \wedge \psi & \text{iff} & M, w \models \varphi \text{ and } M, w \models \psi \\
M, w \models \varphi \vee \psi & \text{iff} & M, w \models \varphi \text{ or } M, w \models \psi \text{ (or both)} \\
M, w \models \varphi \rightarrow \psi & \text{iff} & \text{not } M, w \models \varphi \text{ or } M, w \models \psi \text{ (or both)} \\
M, w \models \varphi \leftrightarrow \psi & \text{iff} & M, w \models \varphi \rightarrow \psi \text{ and } M, w \models \psi \rightarrow \varphi \\
M, w \models \neg \varphi & \text{iff} & \text{not } M, w \models \varphi \\
M, w \models K_i \varphi & \text{iff} & \text{for all } w' \text{ such that } w \sim_i w',\ M, w' \models \varphi \\
M, w \models P_i \varphi & \text{iff} & \text{there exists a } w' \text{ such that } w \sim_i w', \text{ and} \\
& & M, w' \models \varphi
\end{array}
$$

In this definition, the truth conditions for the connectives are exactly like those known from propositional logic. The truth conditions for formulas involving knowledge operators reflect their modal characteristics, namely that having knowledge not only depends on the current situation, but on others as well, as has been expounded above.

For $K_i \varphi$ to be true, agent i must be able to eliminate as possible alternatives to the actual situation all situations in which $\neg \varphi$ holds, thus leaving only situation where φ holds as indistinguishable to him from w.

On the other hand, $P_i \varphi$ will be true in case agent i is *not* able to eliminate all situations in which φ holds, thus leaving at least one state as indistinguishable to him from w on the basis of his current information. An example of a model and some formulas true in it is illustrated in Figure 3.1.3.

As mentioned above, the indistinguishability relations are equivalence relations. With the semantics defined as here, this results in the axiomatic system S5 being sound and complete with respect to the set of all models as defined. That is, a formula can be syntactically proven in S5 if, and only if, it is satisfied in all worlds of all epistemic models, cf. (Fagin et al., 1995).

That the indistinguishability relation is an equivalence relation amounts to it being reflexive, transitive and symmetric, or, equivalently, it being reflexive and euclidean. The axiom T characterizes reflexivity, *4* transitivity and *5* euclideaness, cf. (Blackburn et al., 2001). As reflexivity and euclideaness implies transitivity, this is an easy way of seeing that *4* is implied by T and *5*, as mentioned above.

3.2 Quantified Epistemic Logic

Though propositional epistemic logic allows modeling of a wide range of different epistemic scenarios, the basic ontology is still very limited. Only unstructured propositions are available at the most basic level. This restricts the expressibility in that relations between objects cannot be expressed, and neither can relations between basic propositions based on their internal constitution. Hence, to model knowledge of objects and their properties in a way that preserves structured information, a stronger language is required. To gain expressive power, a more complex syntax as well as semantics is required. For this purpose, a first-order language able to express properties of more structured worlds is introduced. Each world will consist of a domain of individuals and an ascription of properties to these individuals. These worlds will then constitute the state space of epistemic models.

The following exposition of quantified epistemic logic/first-order modal logic is based on (Fagin et al., 1995), (Fitting and Mendelsohn, 1999) and (Hughes and Cresswell, 1996), but includes differences in notation and style of definitions.

3.2.1 First-Order Syntax

Instead of the set *Prop* of unstructured propositions utilized in propositional epistemic logic to define the most basic constituents, the atomic formulas of the language $\mathcal{L}_{\mathsf{QEL}}$ of *first-order quantified epistemic logic* (QEL) are constructed using four sets of constituents:

1. A set of *constant symbols, $CON = \{a, b, c, ...\}$*

2. A set of *variables, $VAR = \{x, y, z, ...\}$*

3. A set of *function symbols, $FUN = \{f, g, h, ...\}$*

4. A set of *relation symbols, $REL = \{P, R, Q, ...\}$*

Both the function symbols and relation symbols[13] may be n-ary,[14] taking n arguments from the set of *terms TER*. The set TER consists of all constant symbols, all variables and where f is an n-ary function symbol and $t_1, ..., t_n$ are terms, $f(t_1, ..., t_n)$ is a term.

The set of *atomic formulas* is defined as follows: where R is an n-ary relation symbol and $t_1, ..., t_n$ are terms, $R(t_1, ..., t_n)$ is an atomic formula. Where '=' is the *identity sign* and t_1, t_2 are terms, $t_1 = t_2$ is an atomic formula.

The atomic formulas play the same role as the propositions did in propositional epistemic logic, but contain more structure and therefore encode more information. To give an example, where in the propositional case '*a loves b*' would be expressed

---

[13]In the present, the terms 'predicate' and 'relation' will be used interchangeably, though relations will in general be assumed to take two or more arguments, but predicates are not taken to be necessarily monadic.

[14]The set of n-ary relation/function symbols will be denoted REL_n and FUN_n, respectively.

by an unstructured proposition p, it may now be expressed by $L(a, b)$, where the binary relation symbol L is interpreted as the relation of 'loving'.

The set of well-formed formulas is defined as follows: any atomic formula is a well-formed formula, and where φ and ψ are well-formed formulas, the following are well-formed formulas:

$$\neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi \mid \varphi \leftrightarrow \psi \mid \forall x \varphi \mid K_i \varphi,$$

where $i \in I$, the set of agents, and $\forall x$ is the *universal quantifier*. The dual of the knowledge operator is defined as in the propositional case, and in the same vein, the dual of $\forall x$, namely $\neg\forall x\neg$, is defined as $\exists x$, the *existential quantifier*. Notions of free and bound variables and sentences are as usual, and can be found in the ensuing chapter.

3.2.2 First-Order Models and Truth Conditions

As mentioned, in the first-order (quantified) case of epistemic logic, worlds are inhabited not by propositions, but by objects and their properties. As a result, the models need to be augmented with a *domain* of individuals, given by a countable set

$$Dom = \{d_1, d_2, ...\}$$

A constant domain for all worlds is chosen, as will be discussed in section 3.2.4 below.

Define the class $\mathcal{C}_{\mathsf{QEL}}$ of *quantified epistemic models* as the set of models M, where M is a quadruple

$$M = \langle W, (\sim_i)_{i \in I}, Dom, \mathcal{I} \rangle$$

such that W and $(\sim_i)_{i \in I}$ are as in the propositional case and $\mathcal{I}$ is a *first-order interpretation* such that:

- Relative to each world, $\mathcal{I}$ assigns to each n-ary relation symbol a set of n-tuples from the domain, i.e.

$$\mathcal{I} : REL_n \times W \longrightarrow Dom^n$$

- Relative to each world, $\mathcal{I}$ assigns to each n-ary function symbol an n-ary function on the domain, i.e.

$$\mathcal{I} : FUN_n \times W \longrightarrow Dom^{Dom^n}$$

  - Where $(t_1, ..., t_n, t') \in \mathcal{I}(f, w)$, it is said that $\mathcal{I}(f(t_1, ..., t_n), w) = \mathcal{I}(t', w)$.

- Relative to each world, $\mathcal{I}$ assigns to each constant an element from the domain, i.e.

$$\mathcal{I} : CON \times W \longrightarrow Dom$$
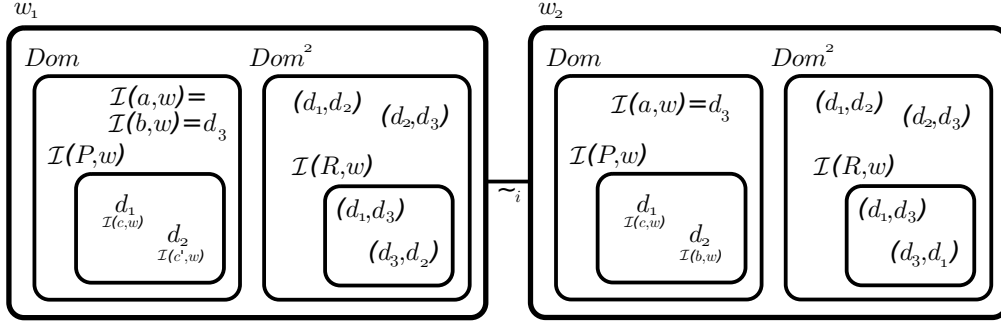
Figure 3.2.1: QEL model with two worlds indistinguishable to i. The domain includes three objects, d_1, d_2 and d_3. A monodic predicate P, and a binary relation R, are represented. Amongst others, the following properties hold: $M, w_1 \models_v (a = b) \wedge \neg P(b) \wedge R(c, a)$ and $M, w_2 \models_v (a \neq b) \wedge P(b) \wedge R(a, c)$. Further, $M, w_1 \models_v K_i(P(c) \wedge R(c, a)) \wedge P_i(a \neq b)$.

Finally, the variables are assigned values world independently by a *valuation, v,* which is a function

$$v : VAR \longrightarrow Dom,$$

assigning to each variable an element of the domain. A valuation v' which assigns to all variables except possibly x the same values as v will be call an *x-variant of v*.

Based on these definitions, the truth conditions may now be defined. In the first-order case, and extra parameter is introduced, namely the valuation. As a result, the satisfaction relation is indexed by this. '$M, w \models_v \varphi$' is read 'in model M at world w, under valuation v, φ is true'. In the below table, some definitions for the basic connectives has been left out, as these are as for the propositional case:

$$
\begin{array}{lll}
M, w \models_v P(t_1, t_2, ..., t_n) & \text{iff} & (d_1, d_2, ..., d_n) \in \mathcal{I}(w, P) \\
M, w \models_v (t_1 = t_2) & \text{iff} & d_1 = d_2 \\
& \text{where} & d_i = \begin{cases} v(t_i) & \text{if } \tau_i \in VAR \\ \mathcal{I}(w, t_i) & \text{if } \tau_i \in CON \end{cases} \\
& \text{for} & i \in \{1, 2, ..., n\} \\
M, w \models_v \varphi \wedge \psi & \text{iff} & M, w \models_v \varphi \text{ and } M, w \models_v \psi \\
M, w \models_v \neg\varphi & \text{iff} & \text{not } M, w \models_v \varphi \\
M, w \models_v K_i\varphi & \text{iff} & \text{for all } w' \text{ such that } w \sim_i w', M, w' \models_v \varphi \\
M, w \models_v P_i\varphi & \text{iff} & \text{there exists a } w' \text{ such that } w \sim_i w', \text{ and} \\
& & M, w' \models_v \varphi \\
M, w \models_v \forall x \varphi(x) & \text{iff} & \text{for all } x\text{-variants } v' \text{ of } v, M, w \models_{v'} \varphi(x) \\
M, w \models_v \exists x \varphi(x) & \text{iff} & \text{for some } x\text{-variant } v' \text{ of } v, M, w \models_{v'} \varphi(x)
\end{array}
$$

The truth values of sentences depend only on models, not on specific valuations, and hence each sentence will in any model have a definite truth value.

An example of a quantified epistemic model is illustrated in Figure 3.2.1.

3.2.3 Philosophical Interpretations: Classes, Concepts and more Indistinguishability

In the propositional case, the most basic entity about which the agents could be uncertain are the unstructured propositions of *Prop*. The situation gains complexity when these propositions gain internal structure. Where in the propositional case, the truth or falsity of a proposition at a given world in some model only depended upon what subset of *Prop* was assigned to that world by the interpretation, the truth of the formulas in the quantified case ranges from at least including an interpretation of a constant and a monadic predicate to more complex cases where interpretations of n constants and a m-adic predicate and $m - n$ existential or universal judgments are required to determine the truth value of the given formula. All these factors can individually be subject to a kind of uncertainty not present in the propositional case. To get an idea of the troubles the agents go trough, the philosophical interpretation of the mentioned constituents of quantified epistemic logic will be presented in the following section.

It should be noted that the exposition of quantified epistemic logic presented here is not an industrial standard, as opposed to the propositional case. The reason for this is that no general standard exists. The formal system is well-known, and presented in (Fagin et al., 1995), but the interpretation presented there is not detailed enough to yield a philosophical basis. The current interpretation is partly based on various texts by Hintikka, e.g. (1962; 1969; 1994; 2007), but no blame should fall on Hintikka if the reader should find the current presentation incoherent.

The system differs in important aspects from the system presented in (Hintikka, 1962), in it having constant domains consisting of the same individuals.[15] Constant domains and the utilized semantics makes cross-identification across epistemic alternatives easy, as two constants (seen as objects by the agents) are the same if their interpretation are the same object in the domain. This is in contrast to (Carlson, 1988), where objects are individuated by *individuating functions*. To exemplify, the variable valuation of (Carlson, 1988) assigns to each variable a partial function from worlds to the domain of each world with certain requirements.

The constants and predicates of quantified epistemic logic need two distinct interpretations depending on whether or not they occur within the scope of knowledge operators. When predicates and constants occurs outside of the scope of operators, these are interpreted as denoting sets of objects (or single objects, in the case of constants), marking some (i.e. the modelers) classification of the objects of the domain, relative to the given world. Thus, outside the scope of operators, predicates denote *properties* of the objects at the world of evaluation, the properties that the modeler sees fit to be denoted relative to the modeling task at hand.

When a predicate (or constant) occurs within the scope of an operator, it represents the agent's *concept*[16] (or *individual concept* in the case of constants) of the

---

[15]The system presented by Hintikka has a completely different semantics, utilizing his *model sets.*

[16]Or, perhaps better, the agents *conception* of the given class, where this conception would give rise to the agent's concept of the class.

property (or object) that the predicate (or constant) denotes outside the scope of the operator, relative to the given world. Before moving on to the important subtleties added by quantifiers, a few notes about these first two aspects are in place.

Predicates and constants outside the scope of operators   In regard to the first, the assignment of properties to the objects of the domain relative to a given world, this assignment is given by the modeler when defining the interpretation of the predicates. Whether this assignment is meant to mirror either objective properties, i.e. as describing natural kinds truly existing in the real world, or as socially constructed properties (or any mixture of the two), makes no difference in the formal process. Hence any application of formal models on real life situations require arguments for the proposed assignment holding true of the real scenario.

Here, no one, *specific* purpose for the models invoked are assumed: they will be used to model semantic competence, but what objects and properties that are relevant will depend on a specific application. As pointed out in (Sokal, 1977) regarding classifications, these tend to be teleological: for some given purpose, some classification may be more productive than others. If ones purpose is the "search for immanent structure in reality" (*ibid.*, p.189), one may aim to supply a taxonomy isomorphic with the given natural system. If one on the other hand aims to analyze a card game, it may be the more pragmatic choice to supply classes such as 'Hearts', 'Spades' and 'Aces', etc. The same goes for the present modeling framework: if one wishes to analyze a given card game, it will typically not make any sense to model properties of the players shoes, whereas the color of the individual cards may be quite interesting.

In the same spirit, constants occurring outside the scope of operators denote objects. Therefore constants outside the scope of operators will be used to denote the real-world objects in Marconi's SLC.

Binarity   One embedded restriction in the current system is its binarity: given the truth conditions stated above, each element either falls into a given class or it does not. This is embedded in the validity of *Tertium non datur*[17],

$$P(x) \vee \neg P(x)$$

stating that for any x either x is P or it is not. In particular, this means that certain structures become problematic to model, namely those the classes of which are thought to grade into one another in a continuous fashion. Here, two cases in point may be Sorites's "pile"[18] and the socially constructed concept of 'species': a

---

[17]Tertium non datur is also known as the Law of the Excluded Middle, or Principle of the Excluded Middle.

[18]Sorites's "pile" is the main character of the *Sorites Paradox*: imagine a pile of sand. If one grain of sand is removed, does the remained still constitute a pile? Yes. As the pile was arbitrary, removing one grain of sand from any pile of sand results in a pile of sand. Iterating the process, the remaining pile of sand includes only one grain, which is not a pile. Contradiction. The

common criterion (often seen as *the* defining criterion) for determining whether two animals are of the same species, is whether their offspring is fertile. Though this may seem like a clear cut criterion it is, as Sorites's "pile", vague: for whether two (fertile) animals produce fertile offspring is a matter of chance, as they can each belong to the outer rim of the given species, and get fertile offspring at one point, but not at another, depending on the exact combination of spermatozoon and oocyte – cf. (Barton et al., 2007, p. 620-622).

What must be noted here is that though it may seem safe to assume non-vague classes of certain objects such as American citizens, elementary particles and red/black socks, in certain other, intuitively clear, cases, a clear cut non-vague classification may be hard if not impossible to obtain – even given expert knowledge of the subject.

Inside the scope of operators When it comes to predicates and constants occurring within the scope of operators, these are not to be interpreted as factual in the same way as when occurring outside such scope. Instead, predicates and constants inside the scope of operators are interpreted as representing the agent in question's subjective concept/conception of the given properties and objects.[19] As such, predicates and constants inside the scope of operators plays the role of the items in the semantic lexicon of Marconi. The details of this is explained in section 5.3.1.

To get a grip of the current use of such concepts, recall that when a formula, say $P(a)$, occurs within the scope of a knowledge operator, like in $K_i P(a)$, the truth of the latter requires that the former is true in all worlds indistinguishable for agent i. Thus, in all the relevant worlds w', it must be the case that $\mathcal{I}(a, w') \in \mathcal{I}(P, w')$. But notice, for two worlds, w and w', both the interpretation of a and that of P may differ from one world to the other.

In the case where for instance the interpretations of P differ between w and w', then if w and w' are indistinguishable for agent i, then the two objectively different properties $\mathcal{I}(P, w)$ and $\mathcal{I}(P, w')$ will also be indistinguishable for i. Hence, these two properties, indistinguishable to the agent, will come under one heading, P, to the agent, and it is this which is meant by a concept or the agents conception of the property. For example, when philosopher Hilary Putnam famously cannot tell elms from beeches, cf. (Putnam, 1975), his concepts of each of the two objectively different classes is the same.

The same applies to constants and individual concepts: if the interpretation of some constant a differs across the worlds indistinguishable to some agent, the agent has a concept of a, but this concept is ambiguous. Perhaps the agent cannot tell the apple on the table from the pear on the table or cannot distinguish between the Chanel purse and the Gucci bag – it's all the same to him.

These cases of ambiguous (individual) concepts are ones in which multiple proper-

---

Sorites Paradox has been used to argue for the *vagueness* of the term pile: the term is vague as no clear boundary can be decided between what constitutes a pile, and what does not.

[19]When occurring within the scope of a knowledge operator, these concepts will retain some factuality, though, due to the reflexivity of the indistinguishability relation.
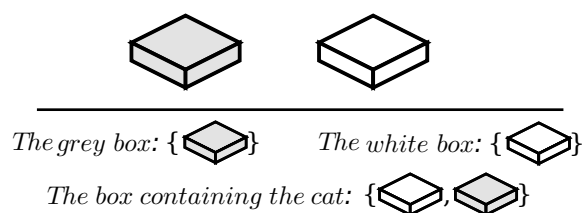
Figure 3.2.2: Two boxes, one of which contains a cat. The cat does not make any sounds, as otherwise one could tell. The white box can be individuated, so can the grey. The box containing the cat cannot. Each would be denoted by constants.

ties (objects) indistinguishable to the agent are thought of as being lumped together under the same mental heading. This does not imply that the agent will perceive for example the two bags as *one object*. It simply means that they are indistinguishable by the headings *Gucci* and *Chanel*. They may still be distinguishable by other headings, like color, position, shape or price. A similar example is illustrated in Figure 3.2.2.

The agent may also be unknowing in a reverse sense, namely by a single thing under two to the agent distinct headings. This may for instance be the case when an agent does not know that the ace of spades and the card marked with an X on the back is the same, or when the agent's concept of 'renates' differs from that of 'cordates'.

Quantifiers As with other formulas, formulas involving quantifiers may occur wholly outside the scope of operators, wholly inside the scope of operators, or, which is a feature that yield a truly novel expressibility of quantified epistemic logic over propositional logic, the scope of quantifiers and operators may be mixed.

Given the above interpretations of formulas not involving quantifiers, the first two cases are rather straightforward. In the case where a quantified formula occur wholly outside the scope of operators (that is, the quantifier occurs outside the scope of operators, and no bound variables occur inside the scope of operators), the role of the formula is that of a normal first-order logic formula. It expresses properties of existence and universality of the objects of the domain and its truth value only depend on the interpretation of predicates and constants in the world of evaluation.

When a quantified formula occur wholly inside the scope of operators, existence or universality is claimed of the agent's concepts occurring within the formula. Certain validities of these types of formulas may occur counter-intuitive, as for instance the validity of

$$K_i \exists x \, (x = a)$$

where a may be any constant. This formula expresses the fact that agent i knows that his concept of a is not empty – there is an object that is a. The validity of this formula is due to the fact that the interpretation is defined as a total function instead of a partial function as is done elsewhere, as for example in *free logics*, cf.

(Garson, 2006).

The validity represents that what the agents in the scenarios modeled are to think about are also assumed to be existing. Hence, if one wished to protest that the formulas of the given type ought not be valid due to obvious counter-examples, e.g. that many have a concept of Sherlock Holmes, well-knowing that no such figure exists, the equally obvious reply would be that an application requiring the modeling of non-existing objects is not what is being presented.

Where the scope of quantifiers and operators are mixed (i.e. when a quantifier is outside the scope of an operator, but there is a bound variable inside the scope of the operator), some very interesting expressibility is gained. Formulas with such mixed quantifier and operator scope will be of crucial importance in relation to the definitions of the referential competence types.

What is made possible by such mixed scopes is the construction of formulas pertaining to attitudes towards specific objects of the actual world. Most notably, it gives the possibility of constructing *wh*-knowledge[20] with the knowledge operators, as introduced in (Hintikka, 1962). Here, the reading of

$$\exists x K_i \left(x = a \right) \tag{3.2.1}$$

is 'there exists some object x of which i knows that x is the object a'. Put more plainly: 'i knows who a is', cf. (Hintikka, 1962, 1984, 1994; Hintikka and Sandu, 1995; Fagin et al., 1995). This interpretation is warranted by the workings of the semantic machinery in that the truth of (3.2.1) ensures that agent i has an unambiguous concept of a. This is due to the fact that the existential quantifier binds the variable x to some object in the domain, say d, at the world of evaluation and subsequently the knowledge operator states that in all worlds indistinguishable to i, the value of the interpretation function of a is d. That is, $\mathcal{I}\left(a, w' \right) = d$ for all w' such that $w \sim_i w'$. Said differently, the interpretation of the constant does not change across the agent's epistemic alternatives. A *de re*[21] formula thus expresses *of the same object* across multiple worlds that it possess some quality – in the case of (3.2.1) that of being identical to a.

In Figure 3.2.2, the illustration of the two boxes is what the agent would perceive, and under the assumption that the agent knows that one of the boxes contains a cat, an epistemic model of this scenario can be illustrated as in Figure 3.2.3. In the figure, the domain consists of two objects, the left box, d_1, and the right box, d_2. In order to speak of the grey box, the white box and the box containing the cat, denote these by constants g, w and c, respectfully. Hence, the illustration shows that $\mathcal{I}(g, w_1) = \mathcal{I}(g, w_2) = d_1$ and $\mathcal{I}(w, w_1) = \mathcal{I}(w, w_2) = d_2$, i.e. that the interpretation of g and w is constant across the two states. Further, it shows that the interpretation of c varies: $\mathcal{I}(c, w_1) = d_1$ and $\mathcal{I}(c, w_2) = d_2$. Hence, it would hold for agent i that

$$\exists x K_i(x = g) \wedge \exists y K_i(y = w) \wedge \neg \exists z K_i(z = c)$$

---

[20]Knowledge of who, what, where, etc., depending on the interpretation of the domain of quantification.

[21]Statements *de re* (as $\exists x K_i \varphi \left(x \right)$) are opposed to statements *de dicto* (as $K_i \exists x \varphi \left(x \right)$). The former are statements *about the thing*, whereas the latter are statements *about the fact*.
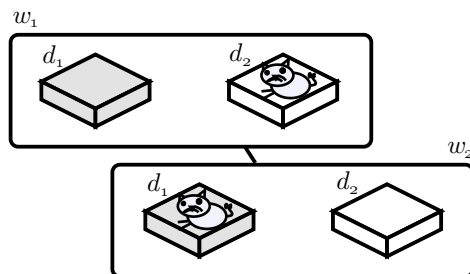
Figure 3.2.3: Illustrated model with two worlds. In each, the grey box is denoted g, the white box w and the box containing the cat, c. The agent can identify the grey and the white boxes, but not the one containing the cat.

i.e. i is able to identify g, the grey box, w, the white box, but not c, the box containing the cat.

To give a further example, the formula

$$\exists x(K_{Putnam}Tree(x) \wedge Elm(x))$$

could be used to express that Putnam knows of some object that it is both a tree and a elm, which would be false. Truthfully, it could be stated that there exists an elm, and of that elm, Putnam can identify it as a tree:

$$\exists x(Elm(x) \wedge K_{Putnam}Tree(x))$$

The type of knowledge towards one particular object as expressed in (3.2.1) is rather strong, and cannot merely be assumed. In particular, though $K_i(a = a)$ is a validity, it may not be concluded by existential generalization that $\exists x K_i\,(x = a)$. This inference is invalid, as will be commented on further in the subsequent section.

Taking *de re* formulas as (3.2.1) to capture the meaning of the notion on 'knowing who' as used in natural language, as it seems done by Hintikka and Fagin et. al, has been strongly criticized in (Boer and Lycan, 1986). In order not to misrepresent the meaning of *de re* statements, the 'knows who' reading of such will not be utilized as the primary one. Instead, such formulas will be interpreted as expressing an *identificatory ability.* Under this interpretation, (3.2.1) states that agent i is able to identify a particular object as the object of his unambiguous concept a. This interpretation also occurs in the works of Hintikka, as when he calls formulas such as (3.2.1) 'identification statements' (1969, p. 161). In the terminology of (Hintikka and Symons, 2007), the interpretation here is *subject-centered* or *perspectival.* A formal definition of individual concepts will be introduced in chapter chapter 5.

Recall that no explicit incorporation of the means of justification are presented in the formal framework, and that the methods by which information is gained is not incorporated in the interpretation of the knowledge operators. This can make the notion of identification seem artificial, as one usually identifies objects by using one or another *method.* A friend is *seen*, the child is *heard*, something is *tasted* as

salt, etc., but seldom anything is identified without using any method for doing so. Hence, identification statements only express that an agent is able to perform the given identification, but there is no concern as to *why* the agent has this ability, nor *how* the identification will be performed.

3.2.4 First-Order Axioms and Inference Rules

The axiom system which is sound and complete with respect to the semantics for quantified epistemic logic presented in the preceding sections, is a proper extension of S5. The axioms of Quantified S5, QS5, includes all the axioms of S5 along with axioms of first-order logic with identity, namely

$$
\begin{array}{ll}
\text{Id:} & t = t \\
\exists\text{Id:} & (c = c) \rightarrow \exists x\,(x = c) \\
\text{UI:} & \forall x \varphi \rightarrow \varphi\,(y/x) \\
\text{SF:} & (t_1 = t_2) \rightarrow (f\,(...,t_1,...) = f\,(...,t_2,...)) \\
\text{PS} & (x = y) \rightarrow (\varphi\,(x) \leftrightarrow \varphi\,(y))
\end{array}
$$

Notice that the *Principle of Substitution*, PS, and *Universal Instantiation*, UI, are restricted to variables. If all terms where allowed, these axioms would lose their validity, and hence result in an unsound system. This is a consequence of including non-rigid constants, as pointed out in (Fagin et al., 1995).

Simply adding first-order axioms to S5 is not sufficient. This is a result of the inherent interaction of quantificational and modal principles in the semantics. In order to handle such, two mixed axioms are required for completeness, namely the *Barcan Formula*, BF, and *Knowledge of Non-identity*, KNI:

$$
\begin{array}{ll}
\text{BF:} & \forall x K_i \varphi\,(x) \rightarrow K_i \forall x \varphi\,(x) \\
\text{KNI:} & (x_1 \neq x_2) \rightarrow K_i\,(x_1 \neq x_2)
\end{array}
$$

Both BF and KNI are included for all agents $i \in I$. The role of KNI is not of particular interest here: its role is primarily technical, and necessary for completeness with respect to the given semantics as the valuations v are world-independent. It should be noticed, though, that KNI too is included in a version restricted to variables again due to the fact that the interpretation of constants is world-dependent, and the unrestricted version would thus not be valid. The opposing principle, *Knowledge of Identicals*,

$$
\text{KI:} \quad (x_1 = x_2) \rightarrow K_i\,(x_1 = x_2)
$$

is deducible from the other axioms, cf. (Fagin et al., 1995).

The Barcan Formula in combination with the remaining axioms allows for the derivation of the *Converse Barcan Formula*,

$$
\text{CBF:} \quad K_i \forall x \varphi\,(x) \rightarrow \forall x K_i \varphi\,(x)
$$

as shown in (Hughes and Cresswell, 1996). Both BF and CBF are valid in constant domain semantics. In fact, these two formulas characterize exactly the constant domain aspect of the first-order modal models, cf. (Fitting and Mendelsohn, 1999). If world-relative domains were assigned by a domain assigning function D such that

$$D : W \longrightarrow \mathcal{P}\left(Dom\right)$$

i.e. the domain assignment function assigned subsets of the domain to each world, but the semantics were otherwise left unchanged, the Barcan Formula would characterize the criteria that if two worlds were related by $\sim_i$, i.e. if $w \sim_i w'$, then their domains were monotonically decreasing in the direction of the relation, i.e. $D\left(w'\right) \subseteq D\left(w\right)$.

The Converse Barcan Formula characterizes the converse criteria, namely that if $w \sim_i w'$, then $D\left(w\right) \subseteq D\left(w'\right)$. Hence, when $\sim_i$ is assumed to be an equivalence relation – and therefore symmetric – it is no wonder that both formulas are valid, as they in conjunction require that if $(w, w') \in \sim_i$, then $D\left(w\right) = D\left(w'\right)$.[22]

The Barcan and Converse Barcan Formulas are hence characteristic of constant domain semantics, a feature of the present logic that has not yet received the attention it deserves. In combination with the total interpretation function, this restriction gives agents a full overview of the world they are knowledgeable of. For the inclusion of the Barcan formulas and, going with them, the choice of constant domain semantics results in the agents "knowing the domain". For example, if some object a exists, then this facts is know, i.e.

$$\exists x \left(a = x\right) \rightarrow K_i \exists x \left(x = a\right)$$

is valid. Further, if the cardinality of the domain is n, then this is known, i.e.

$$\exists x_1 \exists x_2 \exists x_3 .... \exists x_n [(x_1 \neq x_2) \wedge (x_1 \neq x_3) \wedge ... \wedge (x_1 \neq x_n) \wedge (x_2 \neq x_3) \wedge ... \wedge (x_2 \neq x_n) \wedge ... \wedge$$
$$(x_{n-1} \neq x_n)] \rightarrow K_i \exists x_1 \exists x_2 \exists x_3 .... \exists x_n [(x_1 \neq x_2) \wedge (x_1 \neq x_3) \wedge ... \wedge (x_1 \neq x_n) \wedge (x_2 \neq x_3) \wedge ... \wedge$$
$$(x_2 \neq x_n) \wedge ... \wedge (x_{n-1} \neq x_n)]$$

is valid. This is a feature which makes the current logic inapplicable to a large quantity of real-life situations. For example, the cardinality of the domain of voters is not known in many voting situations, where this may have an effect on the strategic decision. In other voting situations, that the domain of voters is common knowledge is an important aspect and can have a large effect on the outcome. In some situations, the feature is not only plausible, but required: imaging, for example, playing cards without knwoing the size of the deck, or playing Nim without knowing the size of the pile.

---

[22]Of course, one could have worlds in a model that were not related by any of the appropriate relations why their domains would not be demanded to be equal, but as pointed out in (Blackburn et al., 2001), then if some world(s) have absolutely no connection to the relevant part of a given model, then the truth values of the formulas of the relevant part of the model will not be changed were the former world(s) removed from the model all together. This feature of modal logic is called the *Sub Model Property*.

Inference Rules   As discussed in section 3.2.3, existential *de re* formulas like

$$\exists x K_i \varphi(x)$$

has a special role in QEL. In such formulas, the variable quantified over is bound to a specific object at the world of evaluation and the formula then "speaks of" that object in the worlds quantified over by the given operator(s). Such formulas, often said to be *quantifying in*, have caused quite some discussion in the literature of quantified modal logic and has notoriously caused Quine to oppose to the project as a whole, cf. (Fitting and Mendelsohn, 1999). The problem is that due to the standard first-order version of the classic inference rule *Existential Generalization*

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

Using this inference rule, counter-intuitive conclusions may be drawn. This may be illustrated using an example of Hintikka's (2007). Assume that a detective, d, is to solve a murder in a to him unknown village. Upon arrival, he is truthfully told that the murderer of the deceased John Doe, j, is the village doctor, v. This induces that

$$K_d M(v, j)$$

Hence, the detective knows that the village doctor murdered John Doe. Using standard existential generalization, it now follows that

$$\exists x K_d M(x, j)$$

This has the interpretation that the detective can identify the murderer. The detective can thus make the arrest asking no further questions. As the detective had no previous knowledge of the village and it's inhabitants, this conclusion is too strong.

   The detective lacks what Hintikka calls the *conclusiveness condition*. In the present terminology, the agent lacks identifying knowledge of the village doctor. Thus, when involving knowledge operators, the rule guarding existential generalization must be changed from the standard version, namely

$$\frac{K_i \varphi(a)}{\exists x K_i \varphi(x)}$$

where $a \in CON$ and φ is an arbitrary formula, to the modified version

$$\frac{K_i \varphi(a) \wedge \exists x K_i (x = a)}{\exists x K_i \varphi(x)}$$

as argued in (Hintikka and Sandu, 1995). The latter version explicitly incorporates the requirement of identifying knowledge regarding a.

   The switch is not warranted only by intuition. The former rule of inference does not preserve truth in semantics with non-rigid constants. Thus, if one wishes to keep

a system of contingent identity, which in turn is required if the agents should not be able to identify all objects of the domain by default, the latter rule of inference must be adopted as the rule of existential generalization. Equivalently, the first-order axiom $\varphi(t) \to \exists x \varphi(x)$, where t is any term, should be restricted to $t \in VAR$, which is exactly what was done when UI above was restricted to variables only.

With possible confusion regarding the rule of existential generalization out of the way, it is now easy to list the rules of inference required for a complete axiomatic system of the logic described. The required rules are Knowledge Generalization and Modus Ponens as presented in section 3.1.3, and *Universal Generalization*, UG:

$$\frac{\varphi \to \psi}{\varphi \to \forall x \psi}$$

where x does not occur free in φ.

The logic QS5 is defined as the smallest set of formulas containing all mentioned axioms which is closed under the mentioned inference rules. This logic is sound and complete with respect to the semantics as defined. This is reported in (Fagin et al., 1995), though for unknown reasons, the proof is not supplied. Nor has such a proof been found in the literature. The result can easily be established using the the *Canonical Class Theorem* proven in the next chapter, and it will be shown there. The logic QS5 and the presented semantics will in the following be referred to simply as quantified epistemic logic, QEL.

4 Many-Sorted Modal Logic

In the present exposition of many-sorted modal logic and it's completeness, the various definitions and proof-techniques are inspired by the works (Blackburn et al., 2001; Fagin et al., 1995; Ebbinghaus et al., 1994; Hughes and Cresswell, 1996; Zarba, 2006). Neither of the works combine the material as is done here. In (Zarba, 2006), basic definitions, syntax and semantics are presented for a non-modal many-sorted logic, but no axiom system is given. Completeness of a sequence calculus for single-sorted first-order logic is found in (Ebbinghaus et al., 1994). (Blackburn et al., 2001) provides basic definitions, syntax, semantics and proof-techniques for the completeness of many propositional modal logics using canonical models. The Canonical Class Theorem is an adaption of their Canonical Model Theorem, see p. 199. (Hughes and Cresswell, 1996) provides an axiomatization and a completeness proof for constant domain semantics with identity, but does not include constants nor function symbols. (Fagin et al., 1995) presents an axiom system for constant domain semantics with identity and non-rigid constants, but does not prove the completeness.

The following introduces a system of modal many-sorted logic with constants, function symbols and identity. The semantics are constant domain semantics and allows for a combination of rigid and non-rigid constants, depending on the sorts in the chosen language. The main proof is a variant of the above mentioned Canonical Model Theorem shown for arbitrary (countably) many sorts and arbitrary (finitely) many operators.

4.1 Syntax

Definition 4.1 ($\mathcal{L}_{n,\overline{\sigma}}$ **Language**)**.** Define a *many-sorted modal language* $\mathcal{L}_{n,\overline{\sigma}}$ to consist of

1. A countable set $\overline{\sigma}$ of *sorts.*

2. An countably infinite set VAR of variables, each assigned a sort $\sigma \in \overline{\sigma}$. The set of σ variables is denoted VAR_σ and is assumed countably infinite for each sort.

3. A possibly empty countable set CON constant symbols, each assigned a sort $\sigma \in \overline{\sigma}$. The set of constants of sort σ is denoted CON_σ. Each set CON_σ is said to be either *rigid* or *non-rigid.*

4. A possibly empty countable set FUN of function symbols, each assigned an *arity* $\sigma_1 \times \cdots \times \sigma_n \longmapsto \sigma$ with $n \geq 1$ and $\sigma_1, ..., \sigma_n, \sigma \in \overline{\sigma}$. The set of function symbols of arity $\sigma_1 \times \cdots \times \sigma_n \longmapsto \sigma$ is denoted $FUN_{\sigma_1 \times \cdots \times \sigma_n \longmapsto \sigma}$.

5. A possibly empty countable set REL of relation symbols, each assigned an arity $\sigma_1 \times \cdots \times \sigma_n$ with $n \geq 1$ and $\sigma_1, ..., \sigma_n \in \overline{\sigma}$. The set of relation symbols of arity $\sigma_1 \times \cdots \times \sigma_n$ is denote $REL_{\sigma_1 \times \cdots \times \sigma_n}$.

6. The identity symbol $=$.

7. A set modal operators K_i, one for each $i \in I = \{1, ..., n\}$.

8. The logical connectives $\neg$ and $\vee$.

9. The universal quantifier $\forall$. ◄

In relation to 3., it should be noted that no notation is introduced for rigidity. Whether a set of constants is assumed rigid or not will be clear from context. Further, it is assumed that a set of constants is rigid if, and only if, it is not non-rigid.

Definition 4.2 (σ-terms). Define the set of $\mathcal{L}_{n,\overline{\sigma}}$ *σ-terms* by the smallest set TER_σ such that

1. $CON_\sigma \subseteq TER_\sigma$

2. $VAR_\sigma \subseteq TER_\sigma$

3. Where f is a function symbols of arity $\sigma_1 \times \cdots \times \sigma_n \longmapsto \sigma$ and t_i is a term of sort σ_i for $i = 1, ..., n$, $f(t_1, ..., t_n)$ is a term of sort σ. ◄

Definition 4.3 (Rigid terms). Define the set of $\mathcal{L}_{n,\overline{\sigma}}$ *rigid terms* as the smallest set which includes

1. All sets CON_σ which are rigid

2. The set of $\mathcal{L}_{n,\overline{\sigma}}$ variables, VAR

Definition 4.4 (Well-Formed Formulas). Define the set of $\mathcal{L}_{n,\overline{\sigma}}$ *well-formed formulas* by

1. Where t_i is a term of sort σ_i for $i = 1, 2$, the expression $(t_1 = t_2)$ is an *atomic $\mathcal{L}_{n,\overline{\sigma}}$ well-formed formula*

2. Where t_i is a term of sort σ_i for $i = 1, ..., n$, and P is a relation symbol of arity $\sigma_1 \times \cdots \times \sigma_n$, the expression $P(t_1, ..., t_n)$ *atomic $\mathcal{L}_{n,\overline{\sigma}}$ well-formed formula*

3. All atomic $\mathcal{L}_{n,\overline{\sigma}}$ well-formed formula are $\mathcal{L}_{n,\overline{\sigma}}$ well-formed formula

4. Where φ and ψ are $\mathcal{L}_{n,\overline{\sigma}}$ well-formed formula, $x \in VAR$ and $i \in I$, the following are $\mathcal{L}_{n,\overline{\sigma}}$ well-formed formulas

$$\neg\varphi | \varphi \wedge \psi | \forall x \varphi | K_i \varphi$$

◀

The remaining logical connectives, the existential quantifier and the dual operator of each K_i, denoted P_i, is defined as usual. Well-formed formulas will be referred to simply as formulas.

Definition 4.5 (**Free Variables, Sentences and Substitution**). The free variables of any term t are the variables of t. For any atomic formula φ, the free variables of φ are the variables of the terms of φ. For formulas φ, ψ the free variables of $\neg\varphi$ and $\varphi \wedge \psi$ respectfully are the free variables of φ and of φ and ψ. For formula φ, the free variables of $K_i\varphi$ are the free variables of φ. Finally, the free variables of $\forall x \varphi$ are the free variables of φ apart from x.

Any variable occurring in φ that is not free in φ is called *bound* in φ, and any formula with no free variables is called a *sentence.*

Where φ is a formula, t a term and x a variable, $\varphi(t/x)$ denotes the result of uniformly substituting t for x in φ. That is, $\varphi(t/x)$ is φ where each free occurrence of x has been replaced by an occurrence of t, under the requirement that no free variables of t becomes bound in the process. ◀

4.2 Semantics

Definition 4.6 ($(n, \overline{\sigma})$-**frame**). Define an $(n, \overline{\sigma})$-*frame* for language $\mathcal{L}_{n,\overline{\sigma}}$ to be a triple $\mathcal{F} = \langle W, (\sim_i)_{i \in I}, Dom \rangle$ where

1. W is a non-empty set of *worlds*

2. $(\sim_i)_{i \in I}$ is a set of binary *accessibility relations* on W such that for each $i \in I$, $\sim_i \subseteq W \times W$

3. *Dom* is a non-empty set, the *domain of quantification.* A partition on *Dom* is assumed such that $Dom = \bigcup_{\sigma \in \overline{\sigma}} Dom_\sigma$ and $Dom_\sigma \cap Dom_{\sigma'} = \emptyset$ for all $\sigma, \sigma' \in \overline{\sigma}$. Each Dom_σ is assumed non-empty. ◀

Definition 4.7 (**Interpretation**). Define an *interpretation* $\mathcal{I}$ for language $\mathcal{L}_{n,\overline{\sigma}}$ as a map where

1. If CON_σ is rigid, then $\mathcal{I}$ assigns to each constant of CON_σ an element of Dom_σ

$$\mathcal{I} : CON_\sigma \longrightarrow Dom_\sigma$$

2. If CON_σ is non-rigid, then $\mathcal{I}$ assigns to each constant of CON_σ an element of Dom_σ relative to each state in W

$$\mathcal{I} : CON_\sigma \times W \longrightarrow Dom_\sigma$$

3. $\mathcal{I}$ assigns to each function symbol of arity $\sigma_1 \times \cdots \times \sigma_n \longmapsto \sigma$ and each element of W a set of $n+1$-tuples of $Dom_{\sigma_1} \times \cdots \times Dom_{\sigma_n} \times Dom_\sigma$ such that each assigned set is a function, i.e.

$$\mathcal{I} : FUN_{\sigma_1 \times \cdots \times \sigma_n \longmapsto \sigma} \times W \longrightarrow Dom_\sigma^{Dom_{\sigma_1} \times \cdots \times Dom_{\sigma_n}}$$

4. $\mathcal{I}$ assigns to each relation symbol of arity $\sigma_1 \times \cdots \times \sigma_n$ and each element of W a set of n-tuples from $Dom_{\sigma_1} \times \cdots \times Dom_{\sigma_n}$, i.e.

$$\mathcal{I} : REL_{\sigma_1 \times \cdots \times \sigma_n} \times W \longrightarrow \mathcal{P}(Dom_{\sigma_1} \times \cdots \times Dom_{\sigma_n})$$

◀

Definition 4.8 ($(n, \overline{\sigma})$**-model**). Define a $(n, \overline{\sigma})$-*model* as an $(n, \overline{\sigma})$-frame augmented with an interpretation, denoted $M = \langle \mathcal{F}, \mathcal{I} \rangle$. Where $M = \langle \mathcal{F}, \mathcal{I} \rangle$, the model M is said to be based on frame $\mathcal{F}$. ◀

Definition 4.9 (**Valuation**). Define a *valuation* to be map v assigning to each variable of sort σ an element of Dom_σ, i.e.

$$v : VAR \longrightarrow Dom$$

such that if $x \in VAR_\sigma$, then $v(x) \in Dom_\sigma$.

Further, define an *x-variant of v* to be a valuation v' such that $v'(y) = v(y)$ for all $y \in VAR/\{x\}$. ◀

Definition 4.10 (**Extension**). Define the *extension of term t at world w under valuation v (in the model specified by the context)*, denoted $[\![t]\!]_v^w$, by the following

1. The extension of a non-rigid constant c is $[\![c]\!]_v^w = \mathcal{I}(c, w)$

2. The extension of a rigid constant c is $[\![c]\!]_v^w = [\![c]\!]_v = \mathcal{I}(c)$

3. The extension of a variable x is $[\![x]\!]_v^w = v(x)$

4. The extension for a term $f(t_1, ..., t_n)$ is $[\![f(t_1, ..., t_n)]\!]_v^w = d$
   such that $([\![t_1]\!]_v^w, ..., [\![t_n]\!]_v^w, d) \in \mathcal{I}(f, w)$

The extension $[\![f(t_1, ..., t_n)]\!]_v^w$ may also be denoted $\mathcal{I}(f(t_1, ...t_n), w)$. ◀

Definition 4.11 (**Satisfaction**). Where M is a model, $w \in W$, $\mathcal{I}$ is an interpretation and v a valuation, denote φ being satisfied at w in M under v by $M, w \vDash_v \varphi$, and define the satisfaction relation recursively as follows:

$$
\begin{array}{lll}
M, w \vDash_v P(t_1, ..., t_n) & \text{iff} & (\llbracket t_1 \rrbracket_v^w, ..., \llbracket t_n \rrbracket_v^w) \in \mathcal{I}(w, P) \\
M, w \vDash_v t_1 = t_2 & \text{iff} & \llbracket t_1 \rrbracket_v^w = \llbracket t_2 \rrbracket_v^w \\
M, w \vDash_v \neg\varphi & \text{iff} & \text{not } M, w \vDash_v \varphi \\
M, w \vDash_v \varphi \wedge \psi & \text{iff} & M, w \vDash_v \varphi \text{ and } M, w \vDash_v \psi \\
M, w \vDash_v \forall x \varphi(x) & \text{iff} & M, w \vDash_{v'} \varphi(x) \text{ for all } x\text{-variants of } v \\
M, w \vDash_v K_i\varphi & \text{iff} & M, w' \vDash_v \varphi \text{ for all } w' \text{ such that } w \sim_i w'
\end{array}
$$

◀

Definition 4.12 (Satisfiability and Validity). A formula φ is satisfiable iff there exists a model M, a state w and a valuation v such that $M, w \vDash_v \varphi$. A formula is said to be valid at a state w in a model M iff $M, w \vDash_v \varphi$ for all valuations v, and this is denoted $M, w \vDash \varphi$. A formula is said to be valid in a model M iff $M, w \vDash \varphi$ for all states w, denoted $M \vDash \varphi$. A formula is said to be valid in a frame $\mathcal{F}$ iff φ is valid in all models based on $\mathcal{F}$, and this is denoted $\mathcal{F} \vDash \varphi$. A formula is said to be valid on a class of frames F iff φ is valid in every frame $\mathcal{F}$ of F, denoted $\mathsf{F} \vDash \varphi$. Finally, a formula φ is said to be valid iff it's valid on the class of all frames, which is denoted $\vDash \varphi$.

Where Γ is a set of formulas, $M, w \vDash_v \Gamma, M, w \vDash \Gamma$ etc., will be used to denote that all formulas of Γ are satisfied in M at w under v, are valid at w in M etc. ◀

Note that where S is a class of models, a model from S is some model M, such that $M \in \mathsf{S}$, but where S is a class of frames, a model from S is a model based on some frame $\mathcal{F}$, where $\mathcal{F} \in \mathsf{S}$.

Definition 4.13 (Consequence Relation). Define the semantic consequence relation for some class of structures S (models or frames) in the following manner: where Γ is some (possibly empty) set of formulas and φ is a single formula, φ is said to be a semantic consequence of Γ over S and write $\Gamma \vDash_\mathsf{S} \varphi$ iff for all models M from S and for all valuations v and all states w in M, if $M, w \vDash_v \Gamma$, then $M, w \vDash_v \varphi$. ◀

The *Principle of Replacement* known from regular first-order logic still holds for the semantics defined above. The principle can be formulated as follows.

Proposition 4.1 (PR). *Let φ be a formula, x, y be variables, M a model, w a state and v a valuation. Then, where v' is a x-variant of v where $v(x) = v'(y)$, $M, w \vDash_v \varphi$ iff $M, w \vDash_{v'} \varphi(y/x)$.*

Proof. Since the only bearing the valuation has on the truth of a formula φ (relative to a state in a model) is the value assigned by the valuation to the free variable of φ, and v' is a x-variant of v (and therefore agree with v on all variables except possibly x) and $\varphi(y/x)$ is exactly like φ except where φ has x free, $\varphi(x/y)$ has free y and $v(x) = v'(y)$, the principle still holds. □

4.3 Normal Axiom Systems

Definition 4.14 ($\mathbf{K}_{n,\bar{\sigma}}$ Axioms). The axioms of $\mathbf{K}_{n,\bar{\sigma}}$ are all substitution instances of validities of propositional logic (that is, where φ is a validity of propositional logic with all propositional variable uniformly replaced by formulas of $\mathcal{L}_{n,\bar{\sigma}}$, φ is an axiom of $\mathbf{K}_{n,\bar{\sigma}}$) and the following axiom schemas:

1. Where φ is any formula of $\mathcal{L}_{n,\bar{\sigma}}$ and t is any rigid term not bound in φ,

$$\forall x \varphi \rightarrow \varphi\left(t/x\right) \tag{$\forall$}$$

2. Where t is any term

$$t = t \tag{Id}$$

3. For all rigid terms t, t', the *Principle of Substitutivity*

$$(t = t') \rightarrow (\varphi\left(t\right) \leftrightarrow \varphi\left(t'\right)) \tag{PS}$$

4. Where c is any constant, *Existences of Identicals*

$$(c = c) \rightarrow \exists x \left(x = c\right) \tag{$\exists$Id}$$

5. For all $i \in I$, axiom K (the *Distribution Axiom*)

$$K_i\left(\varphi \rightarrow \psi\right) \rightarrow \left(K_i \varphi \rightarrow K_i \psi\right), \tag{K}$$

6. For interplay between quantifiers and modal operators, the *Barcan Formula*

$$\forall x K_i \varphi\left(x\right) \rightarrow K_i \forall x \varphi\left(x\right) \tag{BF}$$

7. Where t, t' are rigid terms, *Knowledge of Non-identity*:

$$(t \neq t') \rightarrow K_i\left(t \neq t'\right) \tag{KNI}$$

◀

The axioms $\forall$, PS and KNI must be restricted to rigid terms only as these axioms becomes invalid when non-rigid terms are allowed. See (Fagin et al., 1995, p. 88-90) for a proof with respect to regular first-order modal logic. The proof given there carries over almost without change. The axiom *Dual* $(K_i \varphi \leftrightarrow \neg P_i \neg \varphi)$ is not included as the operator P_i was defined as an abbreviation of $\neg K_i \neg$, why Dual is not needed as an axiom.

Definition 4.15 ($\mathbf{K}_{n,\bar{\sigma}}$ Inference Rules). The inference rules of $\mathbf{K}_{n,\bar{\sigma}}$ are

1. *Modus Ponens*

$$\frac{\varphi, \varphi \rightarrow \psi}{\psi}, \tag{MP}$$

2. *Knowledge Generalization* (KG): if φ is a theorem, then $K_i\varphi$ is a theorem.

3. Where x does not occur free in φ, *Universal Generalization*

$$\frac{\varphi \to \psi}{\varphi \to \forall x\psi} \tag{Gen}$$

◀

Using the two above definitions, the logic $\mathbf{K}_{n,\overline{\sigma}}$ can now be defined.

Definition 4.16 (System $\mathbf{K}_{n,\overline{\sigma}}$). Define the *minimal many-sorted modal logic based on $\mathcal{L}_{n,\overline{\sigma}}$*, denoted $\mathbf{K}_{n,\overline{\sigma}}$, as the smallest set of formulas that includes all $\mathbf{K}_{n,\overline{\sigma}}$ axioms and which is closed under the $\mathbf{K}_{n,\overline{\sigma}}$ inference rules. ◀

Definition 4.17 ($\mathbf{K}_{n,\overline{\sigma}}$ Proof). A $\mathbf{K}_{n,\overline{\sigma}}$ *proof is* a finite sequence of formulas, each of which is either a $\mathbf{K}_{n,\overline{\sigma}}$ axiom or is obtained from one or more earlier formulas of the sequence by the application of one of the above rules of inference. A formula φ is $\mathbf{K}_{n,\overline{\sigma}}$ *provable* iff it is the last item of such a sequence. That a formula φ is $\mathbf{K}_{n,\overline{\sigma}}$ provable is denoted $\vdash_{\mathbf{K}_{n,\overline{\sigma}}} \varphi$. ◀

Definition 4.18 (Normal $(n,\overline{\sigma})$ Modal Logics). Any set Λ of $\mathcal{L}_{n,\overline{\sigma}}$ formulas that includes the axioms of $\mathbf{K}_{n,\overline{\sigma}}$ and which is closed under the above inference rules is called a *normal $(n,\overline{\sigma})$ modal logic*. Where $\varphi \in \Lambda$, φ is a *theorem* of Λ, also denoted $\vdash_\Lambda \varphi$. Further, where Λ_1, Λ_2 are logics and $\Lambda_1 \subseteq \Lambda_2$, the logic Λ_2 is an *extension* of Λ_1. ◀

4.3.1 Three $\mathbf{K}_{n,\overline{\sigma}}$ Theorems

Here is an example of a $\mathbf{K}_{n,\overline{\sigma}}$ proof, and three theorems that will be used in the latter.

Proposition 4.2 (KI). *Where t, t' are rigid terms*

$$\vdash_{\mathbf{K}_{n,\overline{\sigma}}} (t = t') \to K_i (t = t')$$

Proof. First notice that $(t = t') \to (K_i (t = t) \leftrightarrow K_i (t = t'))$ is an instance of PS where $\varphi(t) = K_i (t = t)$. By propositional reasoning, this instance implies (1):

PS	(1)	$\vdash_{\mathbf{K}_{n,\overline{\sigma}}}$	$(t = t') \to (K_i (t = t) \to K_i (t = t'))$
(1) and PC	(2)	$\vdash_{\mathbf{K}_{n,\overline{\sigma}}}$	$K_i (t = t) \to ((t = t') \to K_i (t = t'))$
Id and Nec	(3)	$\vdash_{\mathbf{K}_{n,\overline{\sigma}}}$	$K_i (t = t)$
MP on (2), (3)	(4)	$\vdash_{\mathbf{K}_{n,\overline{\sigma}}}$	$((t = t') \to K_i (t = t'))$

PC in (2) refers to Propositional Calculus. As $(\varphi \to (\chi \to \psi)) \to (\chi \to (\varphi \to \psi))$ is a theorem of PC, it is also a theorem of $\mathbf{K}_{n,\overline{\sigma}}$. □

The $\mathbf{K}_{n,\bar{\sigma}}$ theorem

$$(t = t') \rightarrow K_i \, (t = t') \tag{KI}$$

is called *Knowledge of Identity*, following Fagin et al. (1995).

The following principle will later be referred to as *K-distribution* for reasons which are hopefully obvious.

Proposition 4.3 (K-Distribution).

$$\vdash_{\mathbf{K}_{n,\bar{\sigma}}} K_i \, (\varphi_1 \wedge ... \wedge \varphi_n) \leftrightarrow (K_i\varphi_1 \wedge ... \wedge K_i\varphi_n)$$

Proof. The proof is omitted as it is long and completely analogous to the proof in standard modal logic. The case for $n = 2$ can be found in (Hughes and Cresswell, 1996, p. 28) ☐

The following is an unstrict $\mathbf{K}_{n,\bar{\sigma}}$ proof of *Derived Rule 1* from (Hughes and Cresswell, 1996):

Proposition 4.4 (Derived Rule 1).

$$\text{if} \vdash_{\mathbf{K}_{n,\bar{\sigma}}} \varphi \rightarrow \psi$$
$$\text{then} \vdash_{\mathbf{K}_{n,\bar{\sigma}}} K_i\varphi \rightarrow K_i\psi$$

Proof. Assume that $\vdash_{\mathbf{K}_{n,\bar{\sigma}}} \varphi \rightarrow \psi$. Then, by KG, $\vdash_{\mathbf{K}_{n,\bar{\sigma}}} K_i \, (\varphi \rightarrow \psi)$. So, using the appropriate version of K, $\vdash_{\mathbf{K}_{n,\bar{\sigma}}} K_i\varphi \rightarrow K_i\psi$ is obtained. ☐

4.3.2 Deducibility and Consistency

Before defining soundness and completeness, a definition of Λ-consistency, where Λ is a logic, is needed. To define this notion, first define Λ-*deducibility*.

Definition 4.19 (Λ-deducibility). Where $\Gamma \cup \{\varphi\}$ is a set of formulas, φ is single formula and Λ is logic, φ *is Λ-deducible from* Γ if, for some subset $\{\varphi_1, ..., \varphi_n\}$ of Γ,

$$\vdash_\Lambda (\varphi_1 \wedge ... \wedge \varphi_n) \rightarrow \varphi.$$

When φ is Λ-deducible from Γ, this is written $\Gamma \vdash_\Lambda \varphi$, and when φ is not Λ-deducible from Γ, this is denoted $\Gamma \nvdash_\Lambda \varphi$. ◄

Definition 4.20 (Λ-consistency). Where Γ is a set of formulas and φ is any single formula, Γ is Λ-*consistent* if $\Gamma \nvdash_\Lambda \varphi \wedge \neg\varphi$, and Λ-*inconsistent* otherwise. Further, a formula φ is Λ-consistent if $\{\varphi\}$ is Λ-consistent and Λ-inconsistent otherwise. ◄

4.4 Soundness

Definition 4.21 (Soundness). Let S be a class of structures (models or frames) and let Λ be a logic, then Λ is *sound with respect to* S iff for all formulas φ, if $\vdash_\Lambda \varphi$, then $\vDash_\mathsf{S} \varphi$. ◀

If Λ is sound with respect to S, S is said to be a class of frames/models for Λ.
It is now shown that $\mathbf{K}_{n,\overline{\sigma}}$ is sound with respect to the class of all $(n,\overline{\sigma})$-frames, $\mathsf{F}_{n,\overline{\sigma}}$, by showing that 1) all axioms of $\mathbf{K}_{n,\overline{\sigma}}$ are valid in $\mathsf{F}_{n,\overline{\sigma}}$ and 2) the inference rules of $\mathbf{K}_{n,\overline{\sigma}}$ preserve truth in $\mathsf{F}_{n,\overline{\sigma}}$, that is, in any structure $s \in \mathsf{F}_{n,\overline{\sigma}}$, if the hypothesis of a inference rule is satisfied in s, then so is the conclusion.

Lemma 4.1 (Axiom Validity). *All $\mathbf{K}_{n,\overline{\sigma}}$ axioms are valid on $\mathsf{F}_{n,\overline{\sigma}}$*

Proof. Initially note that the validity of the propositional axioms only depend on the semantics for the logical connectives, and since these are completely standard, all propositional tautologies are indeed valid in $(n,\overline{\sigma})$-frames.

$\forall$**:** To see that $\forall$ is valid, assume that φ is any formula of $\mathcal{L}_{n,\overline{\sigma}}$, t is a rigid term and that $M, w \vDash_v \forall x \varphi$ for some model M based on a frame $\mathcal{F} \in \mathsf{F}_{n,\overline{\sigma}}$, state w and valuation v. If the rigid term t is a variable, y, the proof is short. Let v' be an x-variant of v such that $v(x) = v'(y)$. Then, by the semantics for the $\forall$ quantifier, $M, w \vDash_{v'} \varphi$. Thus, by PR, $M, w \vDash_v \varphi(y/x)$. Since $\mathcal{F}, M, w$ and v where arbitrary, this holds for any $\mathcal{F} \in \mathsf{F}_{n,\overline{\sigma}}$, and thus $\forall$ is valid over the class of all $(n,\overline{\sigma})$-frames.

Where the rigid term t is a rigid constant c, the Principle of Replacement from Proposition 4.1 cannot be used, and the proof is on the length of φ. For the atomic cases, assume first that φ is $(x = t')$ and $M, w \vDash_v \forall x \varphi$. Then for all valuations, $[\![x]\!]_v = [\![t']\!]_v^w$. Hence, for all $d \in Dom$, $d = [\![t]\!]_v^w$. Thus $[\![c]\!]_v = [\![t']\!]_v^w$. So $M, w \models_v (c = t')$. Assume second that φ is $P(x, t_2, ..., t_n)$ for some predicate symbol P and terms of appropriate sorts and that $M, w \models_v \forall x P(x, t_2, ..., t_n)$. Then for all valuations, $([\![x]\!]_v, [\![t_2]\!]_v^w, ..., [\![t_n]\!]_v^w) \in \mathcal{I}(P, w)$. So, for all $d \in Dom$, $(d, [\![t_2]\!]_v^w, ..., [\![t_n]\!]_v^w) \in \mathcal{I}(P, w)$. As $[\![c]\!]_v \in Dom$, $([\![c]\!]_v, [\![t_2]\!]_v^w, ..., [\![t_n]\!]_v^w) \in \mathcal{I}(P, w)$. Hence $M, w \models_v P(c, t_2, ..., t_n)$. For φ being $\neg\psi$, assuming $M, w \models_v \forall x \neg\psi$ and $M, w \models_v \neg\psi(c/x)$ leads to a contradiction, as the latter requires that some $d \in Dom$ satisfies ψ, contrary to the first conjunct. Where φ is $\psi_1 \wedge \psi_2$ and $\forall$ holds for ψ_1 and ψ_2, $\forall$ holds for φ as $\forall x(\psi_1 \wedge \psi_2) \leftrightarrow (\forall x \psi_1 \wedge \forall x \psi_2)$ is first-order theorem. Finally, where φ is $K_i\psi$, $M, w \models_v \forall x K_i \psi$ implies that $M, w \models_v K_i\psi(c/x)$ as c is a rigid constant and does therefore not change extension across states.

Id: The validity of Id follows immediately from the semantics for $=$, as defined above.

PS: To see that PS is valid, let $\mathcal{F}, M, w$ and v be as above, let t, t' be rigid terms, and let $M, w \vDash_v (t = t')$ and $M, w \vDash_v \varphi(t)$. Since $M, w \vDash_v (t = t')$, it follows

from the semantics of $=$ that $[\![t]\!]_v = [\![t']\!]_v$. Where φ is an atomic formula, $M, w \vDash_v \varphi(t')$ follows directly. Assuming PS holds for ψ, let φ be $\neg\psi$, and assume that $M, w \vDash_v (t = t')$ and $M, w \vDash_v \neg\psi(t')$. Since PS holds for ψ, and $M, w \vDash_v (t = t')$, $M, w \vDash_v (\psi(t) \leftrightarrow \psi(t'))$. Finally, since assuming $M, w \vDash_v \neg\psi(t)$, $M, w \vDash_v (\psi(t) \leftrightarrow \psi(t'))$ and $M, w \vDash_v \psi(t')$ leads to a contradiction, it is concluded that $M, w \vDash_v \neg\psi(t')$, so PS holds for φ as well. Where $\varphi = \psi_1 \wedge \psi_2$, and PS holds for both ψ_1 and ψ_2, it obviously holds for φ. Finally, where $\varphi(t) = K_i\psi(t)$ and PS holds for ψ, assume that $M, w \models_v (t = t')$. Then $[\![t]\!]_v = [\![t']\!]_v$. Assume $M, w \models_v K_i\psi(t)$. Then for all $w' : w \sim_i w'$, $M, w' \models_v \psi(t)$. As PS holds for ψ and t' is rigid, $M, w' \models_v \psi(t')$. Hence $M, w \models_v K_i\psi(t')$.

$\exists$**Id:** To see that $\exists$Id is valid, notice that all non-rigid constants are being assigned an extension in each w by $\mathcal{I}$, and that all rigid constants are assign and extension relative to the model. Further, the semantics for the $\exists$ quantifier is defined over all possible valuations. Thus, for every constant c, there will be some appropriate valuation v, such that $v(x) = [\![c]\!]_v^w$ for non-rigid constants and $v(x) = [\![c]\!]_v$. Hence, $M, w \vDash_v \exists x (x = c)$.

K: Assume for a contradiction that $M, w \vDash_v \neg (K_i(\varphi \to \psi) \to (K_i\varphi \to K_i\psi))$. Then $M, w \vDash_v K_i(\varphi \to \psi) \wedge K_i\varphi \wedge \neg K_i\psi$. Thus, for all $w' : w \sim_i w'$, $M, w' \vDash_v (\varphi \to \psi) \wedge \varphi$. Hence, by the semantics of the implication, also $M, w' \vDash_v \psi$. But since $M, w \vDash_v \neg K_i\psi$, it follows that $M, w \vDash_v P_i\neg\psi$ so for some $w'' : w \sim_i w''$, $M, w'' \vDash_v \neg\psi$. But then $M, w'' \vDash_v \neg\psi \wedge \psi$, which is impossible.

BF: Assume for a contradiction that $M, w \vDash_v \neg (\forall x K_i\varphi(x) \to K_i\forall x\varphi(x))$. Then $M, w \vDash_v \forall x K_i\varphi(x)$ and $M, w \vDash_v \neg K_i\forall x\varphi(x)$. By the second conjunct, there exists w' such that $w' : w \sim_i w'$ and $M, w' \vDash_v \neg\forall x\varphi(x)$. So, for some x-variant v' of v, $M, w' \vDash_{v'} \neg\varphi(x)$. From the first conjunct, it follows that for all x-variants v'' of v, including v', $M, w \vDash_{v''} K_i\varphi(x)$. Hence, in all w'' such that $w'' : w \sim_i w''$, $M, w'' \vDash_{v''} \varphi(x)$. As these w'' include w', this is absurd.

KNI: Assume that $M, w \vDash_v (t \neq t')$ for two rigid terms t, t'. Then $[\![t]\!]_v \neq [\![t']\!]_v$. As the extension of rigid terms is state-independent, this goes for all $w' \in W$, especially all w' such that $w' : w \sim_i w'$. So $M, w \vDash_v K_i(t \neq t')$.

$\square$

Lemma 4.2 (Validity Preservation). *The $\mathbf{K}_{n,\overline{\sigma}}$ inference rules preserve validity in $\mathsf{F}_{n,\overline{\sigma}}$*

Proof. It is shown that each of the three $\mathbf{K}_{n,\overline{\sigma}}$ inference rules preserve validity:

MP: Assume that both $\varphi \to \psi$ and φ are valid in $\mathsf{F}_{n,\overline{\sigma}}$. Then in all models based on any frame in $\mathsf{F}_{n,\overline{\sigma}}$, in all worlds, under all valuation, both are true. But then, by the semantics of the implication, so is ψ. Hence ψ is valid as well.

KG: Assume that φ is valid in $\mathsf{F}_{\mathsf{n},\overline{\sigma}}$. Then in all models based on any frame in $\mathsf{F}_{\mathsf{n},\overline{\sigma}}$, worlds and valuations $M, w \vDash_v \varphi$. Now let M, w and v be arbitrary. Then $M, w \vDash_v K_i\varphi$ as for all w' such that $w' : w \sim_i w'$, $M, w' \vDash_v \varphi$. So $K_i\varphi$ is valid too.

Gen: Assume for some $M \in \mathsf{F}_{\mathsf{n},\overline{\sigma}}$ that φ is a formula in which x does not occur free and ψ is any formula and argue by contraposition: assume that $M, w \vDash_v \neg(\varphi \to \forall x\psi)$, that is, that $(\varphi \to \forall x\psi)$ is not valid. Then $M, w \vDash_v \varphi$ and $M, w \vDash_v \neg\forall x\psi$ so $M, w \vDash_{v'} \neg\psi$ for a suitable x-variant v' of v. But as x did not occur free in φ, $M, w \vDash_{v'} \varphi$. So $(\varphi \to \psi)$ wasn't valid either.

$\square$

Theorem 4.1 (**Soundness**). *Since all axioms of $\boldsymbol{K}_{n,\overline{\sigma}}$ are valid in $\mathsf{F}_{\mathsf{n},\overline{\sigma}}$, the class of all $(n, \overline{\sigma})$-frames, and the rules of inference all preserve truth with respect to this class of frames, it is concluded that $\boldsymbol{K}_{n,\overline{\sigma}}$ is sound with respect to $\mathsf{F}_{\mathsf{n},\overline{\sigma}}$, that is, if $\vdash_{\boldsymbol{K}_{n,\overline{\sigma}}} \varphi$, then $\vDash_{\mathsf{F}_{\mathsf{n},\overline{\sigma}}} \varphi$.*

4.5 Completeness

Where Λ is a logic and S a class of structures, it is first shown that any Λ-consistent set of formulas is satisfiable on some structure $s \in \mathsf{S}$ if and only if Λ is complete with respect to the class S. This is the content of the proposition IFF below. Then the canonical model for logic Λ and Λ-consistent set of formulas Ω, M_Ω^Λ is defined. Most of the present section concerns the canonical models and related lemmas. The final lemma, the Truth Lemma, states that any Λ-consistent set of formulas Ω is satisfied in the model M_Ω^Λ. Defining the class of canonical model for Λ, M^Λ, to be the set of M_Ω^Λ for any Λ-consistent set Ω yields the conclusion that Λ is complete with respect to M^Λ. This is the content of the main result, the Canonical Class Theorem. From this it can be concluded that if $\mathsf{M}^\Lambda \subseteq \mathsf{S}$, then Λ is complete with respect to the class S. This, in turn, can be used to provide completeness proofs for many-sorted variants of well-known propositional systems.

Definition 4.22 (Completeness). Let S be a class of frames (or models) and let Λ be a logic. Λ is *(strongly) complete* with respect to S if, for any set of formulas $\Gamma \cup \{\varphi\}$, if $\Gamma \vDash_\mathsf{S} \varphi$, then $\Gamma \vdash_\Lambda \varphi$. ◀

Using this definition of completeness, the following result can be obtained, which allowing a proof of completeness using canonical models:

Proposition 4.5 (IFF). *A logic Λ is (strongly) complete with respect to a class of structures S iff every Λ-consistent set of formulas is satisfiable on some structure $s \in \mathsf{S}$.*

Proof. Left to right: Assume that Λ is complete with respect to S. Then every Λ-consistent set is satisfiable on some $s \in \mathsf{S}$. For suppose $\Gamma \cup \{\varphi\}$ is Λ-consistent but

not satisfiable on any $s \in \mathsf{S}$. Then $\Gamma \vDash_\mathsf{S} \neg\varphi$, and so, by completeness, $\Gamma \vdash_\Lambda \neg\varphi$. But then $\Gamma \cup \{\varphi\}$ is Λ-inconsistent contrary to assumption.

Right to left: Assume that Λ is not complete with respect to S. Then, for some set of formulas $\Gamma \cup \{\varphi\}$, $\Gamma \vDash_\mathsf{S} \varphi$ and $\Gamma \nvdash_\Lambda \varphi$. By the second conjunct it follows that $\Gamma \cup \{\neg\varphi\}$ is Λ-consistent, but, by the first, not satisfiable on any structure in S, contrary to assumption. $\qquad\square$

4.5.1 States of the Canonical Model

Using canonical models it will be shown that any Λ-consistent set of $\mathcal{L}_{n,\overline{\sigma}}$ formulas can be satisfied. When constructing the canonical model M_Ω^Λ for some such Λ-consistent set of formulas Ω, the statespace W_Ω^Λ is going to consist of all maximal Λ-consistent sets of formulas satisfying the $\forall$-property. These two notions are defined in this section, where also the basic lemmas are proven.

Definition 4.23 (Maximal Λ-consistent). Where Γ is a set of formulas and Λ is a logic, Γ is said to be *maximal Λ-consistent* iff Γ is Λ-consistent and any proper extension of Γ is Λ-inconsistent. $\qquad\blacktriangleleft$

Where Γ is maximal Λ-consistent, Γ is said to be a a Λ-maximal consistent set, a Λ-MCS. It is further required that each Λ-MCS has the $\forall$-property[1]:

Definition 4.24 ($\forall$-property). Where Γ is a set of $\mathcal{L}_{n,\overline{\sigma}}$ formulas, Γ *has the $\forall$-property* iff for every $\mathcal{L}_{n,\overline{\sigma}}$ formula φ and every variable x, there is some variable y such that $(\varphi\,(y/x) \to \forall x\varphi) \in \Gamma$. $\qquad\blacktriangleleft$

The point of requiring the Λ-MCSs to have the $\forall$-property is that whenever $\forall x\varphi \notin \Gamma$ for some Λ-consistent set with the $\forall$-property, then there must be some witness y such that $\varphi\,(y/x) \notin \Gamma$. For since Γ has the $\forall$-property, $\varphi\,(y/x) \to \forall x\varphi \in \Gamma$ for some y, so if for all y, $\varphi\,(y/x) \in \Gamma$, then $\forall x\varphi \in \Gamma$, contradicting the assumption of Γ's consistency.

Proposition 4.6 ($\forall$-preservation). *If a set of formulas Γ from language $\mathcal{L}_{n,\overline{\sigma}}$ has the $\forall$-property, then for all sets of formulas Δ from $\mathcal{L}_{n,\overline{\sigma}}$ such that $\Gamma \subseteq \Delta$, Δ has the $\forall$-property.*

In order to ensure that any arbitrary set of formulas from $\mathcal{L}_{n,\overline{\sigma}}$ can be extended to a Λ-MCS with the $\forall$-property, there is, in certain cases,[2] need for more variables than those contained in $\mathcal{L}_{n,\overline{\sigma}}$. An augmented language is defined as follows.

Definition 4.25 ($\mathcal{L}_{n,\overline{\sigma}}^+$ Language). Where $\mathcal{L}_{n,\overline{\sigma}}$ is a many-sorted modal language for n agents, let $\mathcal{L}_{n,\overline{\sigma}}^+$ be just like $\mathcal{L}_{n,\overline{\sigma}}$, except that added to $\mathcal{L}_{n,\overline{\sigma}}^+$ is infinitely many *new* variables of each sort $\sigma \in \overline{\sigma}$. Otherwise the two languages are the same, and in particular all well-formed formulas of $\mathcal{L}_{n,\overline{\sigma}}$ are also well-formed formulas of $\mathcal{L}_{n,\overline{\sigma}}^+$. The set of $\mathcal{L}_{n,\overline{\sigma}}^+$ variables is denoted VAR^+. $\qquad\blacktriangleleft$

---

[1]This property is also referred to in the literature as "containing witnesses", "the omega condition" and "saturation".

[2]See (Hughes and Cresswell, 1996, p. 257) for a nice example for standard first-order modal logic.

Three lemmas are proven demonstrating facts regarding sets of formulas needed for the construction of the canonical models.

Proposition 4.7 (Properties of Λ-MCS). *If Λ is a logic and Γ is a Λ-MCS, then*

$$i) \quad \Lambda \subseteq \Gamma$$
$$ii) \quad \textit{for all formulas } \varphi, \textit{ either } \varphi \in \Gamma \textit{ or } \neg\varphi \in \Gamma$$
$$iii) \quad \textit{for all formulas } \varphi, \psi : \varphi \wedge \psi \in \Gamma \textit{ iff } \varphi \in \Gamma \textit{ and } \psi \in \Gamma$$
$$iv) \quad \Gamma \textit{ is closed under Modus Ponens}$$

Proof. i) If Λ is not in Γ, then for some $\varphi \in \Lambda$, $\varphi \notin \Gamma$ so $\neg\varphi \in \Gamma$ as Γ is maximal, so Γ would be Λ-inconsistent contrary to assumption.

ii) If there was some φ such that neither φ or $\neg\varphi$ is in Γ, one of them could be added not yielding an inconsistent set, so Γ was not maximal.

iii) Assume $\varphi \wedge \psi \in \Gamma$ but neither φ nor ψ is in Γ. Then, by *ii)*, $\neg\varphi$ and $\neg\psi$ is in Γ. But then $\{\neg\varphi, \varphi \wedge \psi\} \subseteq \Gamma$ making Γ Λ-inconsistent, as $\neg((\varphi \wedge \psi) \wedge \neg\varphi)$ is a PC theorems and thus in Γ by *i)*. Now, if both φ and ψ are in Γ, but $\neg(\varphi \wedge \psi)$ also is in Γ, the again Γ is Λ-inconsistent by *i)*.

iv) Assume Γ is not closed under MP. Then, for some φ, ψ, $\{\varphi, (\varphi \to \psi), \neg\psi\} \subseteq \Gamma$. So by two applications of *iii)* $(\varphi \wedge (\varphi \to \psi) \wedge \neg\psi) \in \Gamma$. But then Γ is inconsistent, as $\neg(\varphi \wedge (\varphi \to \psi) \wedge \neg\psi) \in \Lambda$. $\qquad \square$

Lemma 4.3 (Lindenbaum's Lemma). *If Γ is a Λ-consistent set of formulas, then there exists a Λ-MCS Γ^+ such that $\Gamma \subseteq \Gamma^+$.*

Proof. Suppose Γ is a Λ-consistent set of formulas from $\mathcal{L}_{n,\bar{\sigma}}$, and let $\varphi_0, \varphi_1, \varphi_2, \ldots$ be an enumeration of the well-formed formulas of $\mathcal{L}_{n,\bar{\sigma}}$. Then define a series of sets Γ_n as follows:

$$\Gamma_0 := \Gamma$$
$$\Gamma_{n+1} := \begin{cases} \Gamma_n \cup \{\varphi_n\} & \text{if this is } \Lambda\text{-consistent} \\ \Gamma_n \cup \{\neg\varphi_n\} & \text{otherwise} \end{cases}$$

and finally

$$\Gamma^+ := \bigcup_{n \geq 0} \Gamma_n$$

Γ^+ is the required Λ-MCS. First, $\Gamma \subseteq \Gamma^+$ by the definition of Γ^+. Secondly, Γ_0 is Λ-consistent by assumption. Where Γ_n is Λ-consistent, and Γ_{n+1} assumed Λ-inconsistent, a contradiction is reached since this means that both $\Gamma_n \cup \{\varphi_{n+1}\}$ and $\Gamma_n \cup \{\neg\varphi_{n+1}\}$ are inconsistent, and thus, there is some set of formulas $\psi_1, \ldots, \psi_m \in \Gamma_n$ such that $\vdash_\Lambda (\psi_1 \wedge \ldots \wedge \psi_m) \to \neg\varphi_{n+1}$ and some set of formulas $\psi'_1, \ldots, \psi'_k \in \Gamma_n$ such that $\vdash_\Lambda (\psi'_1 \wedge \ldots \wedge \psi'_k) \to \varphi_{n+1}$. But as $\vdash_\Lambda ((\psi_1 \wedge \ldots \wedge \psi_m) \wedge (\psi'_1 \wedge \ldots \wedge \psi'_k)) \to (\varphi_{n+1} \wedge \neg\varphi_{n+1})$, it follows by PC that $\vdash_\Lambda \neg(\psi_1 \wedge \ldots \wedge \psi_m \wedge \psi'_1 \wedge \ldots \wedge \psi'_k)$. But since $\psi_1, \ldots, \psi_m, \psi'_1, \ldots, \psi'_k \in \Gamma_n$ and $\Lambda \subseteq \Gamma$ (Property of Λ-MCSs) it follows that both $\neg(\psi_1 \wedge \ldots \wedge \psi_m \wedge \psi'_1 \wedge \ldots \wedge \psi'_k) \in \Gamma_n$ and $(\psi_1 \wedge \ldots \wedge \psi_m \wedge \psi'_1 \wedge \ldots \wedge \psi'_k) \in \Gamma_n$ so Γ_n

is Λ-inconsistent, contrary to assumption. Therefore Γ_{n+1} is Λ-consistent, and thus Γ_m is Λ-consistent for all $m \geq 0$. Thirdly, Γ^+ is Λ-consistent, since if it was not, some finite subset of Γ^+ would have to be Λ-inconsistent, but since every finite subset of Γ^+ is a subset of some Γ_n and Γ_n is Λ-consistent for all n, Γ^+ cannot be Λ-inconsistent. Finally, Γ^+ is a Λ-MCS. Consider any formula φ_n. Either $\varphi_n \in \Gamma_n$ or else $\neg\varphi_n \in \Gamma_n$. Since $\Gamma_n \subseteq \Gamma^+$, either $\varphi_n \in \Gamma^+$ or $\neg\varphi_n \in \Gamma^+$ so any proper extension of Γ^+ is Λ-inconsistent, and thus Γ^+ is a Λ-MCS. $\qquad\square$

Lemma 4.4 (**Saturation**). *If Γ is a Λ-consistent set of formulas from $\mathcal{L}_{n,\overline{\sigma}}$, then there exists a Λ-consistent set of formulas $\Gamma_\forall$ of $\mathcal{L}_{n,\overline{\sigma}}^+$ that has the $\forall$-property and $\Gamma \subseteq \Gamma_\forall$.*

Proof. Assume an enumeration of all formulas of the form $\forall x\varphi$ for any formula φ and of all variables from $\mathcal{L}_{n,\overline{\sigma}}^+$ is given. Define a series of sets Δ_n thus:

$$\Delta_0 := \Gamma$$
$$\Delta_{n+1} := \Delta_n \cup \{\varphi\,(y/x) \to \forall x\varphi\}$$

where $\forall x\varphi$ is the $n+1$th formula in the enumeration and y is the first variable not in Δ_n or φ. Now, since $\Delta_0 = \Gamma \subseteq \mathcal{L}_{n,\overline{\sigma}}$, and Δ_n has been formed by adding only n new formulas, there will still be infinitely many more variables left from $\mathcal{L}_{n,\overline{\sigma}}^+$ to provide the first variable not used, y. Since Δ_0 is assumed Λ-consistent, it is now shown that where Δ_n is Λ-consistent, so is Δ_{n+1}. So assume that Δ_n is Λ-consistent, but that Δ_{n+1} is not. Then there must be formulas $\varphi_1, ..., \varphi_n \in \Gamma_n$ such that

$$\vdash_\Lambda (\varphi_1 \wedge ... \wedge \varphi_n) \to \varphi\,(y/x) \tag{$*$}$$

and

$$\vdash_\Lambda (\varphi_1 \wedge ... \wedge \varphi_n) \to \neg\forall x\varphi$$

Now, since y does not occur free in $(\varphi_1 \wedge ... \wedge \varphi_n)$ since these are from Γ_n, from $(*)$ and Gen it follows that that

$$\vdash_\Lambda (\varphi_1 \wedge ... \wedge \varphi_n) \to \forall y\varphi\,(y/x)$$

but since y does not occur free in φ, $\forall y\varphi\,(y/x)$ is, by standard first-order reasoning[3], equivalent to $\forall x\varphi$. This in turn means that

$$\vdash_\Lambda (\varphi_1 \wedge ... \wedge \varphi_n) \to \forall x\varphi$$

and thus

$$\Gamma_n \vdash_\Lambda \forall x\varphi \wedge \neg\forall x\varphi$$

and Γ_n is thus Λ-inconsistent contrary to assumption, and it is concluded that Γ_{n+1} is consistent. So Γ_m is Λ-consistent for all $m \geq 0$.
Finally, let

$$\Delta^+ = \bigcup_{n \geq 0} \Delta_n$$

By the same reasoning as in the proof for Lindenbaum's Lemma, Δ^+ is consistent, and it has the $\forall$-property by the above construction. $\qquad\square$

---

[3] cf. (Hughes and Cresswell, 1996, p. 242, 258).

4.5.2 Canonical Models

When defining the canonical model for logic Λ and formula set Ω, there is a subtlety regarding identity statements that needs to be taken care of. Simply defining the set of states for a canonical model as the set of all Λ-MCSs as is done in first-order modal logic without identity will not do, since it is needed that the same identity statements between rigid terms to be true in all states of the model. To ensure this, attention is restricted to a *connected* part of the canonical model containing a world which includes Λ.

Definition 4.26 ($\sim$-connected). For some frame $\mathcal{F} = \langle W, (\sim_i)_{i \in I}, Dom \rangle$, let $w, w' \in W$. w and w' are said to be $\sim$-*connected* iff there exists some $w_0, w_1, ..., w_n \in W$ such that $w_0 = w$, $w_n = w'$ and for all $0 \leq k \leq n$ there is some $i \in I$ such that $w_k \sim_i w_{k+1}$. The $\sim_w$-connected sub-frame of $\mathcal{F}$ is a frame $\mathcal{F}_w = \langle W_w, (\sim'_i)_{i \in I}, Dom \rangle$ such that $W_w \subseteq W$ consists of all worlds $\sim$-connected to w and $\sim'_i$ is identical to $\sim_i$ restricted to W_w. ◀

That rigid term identity statements are invariant over states of a $\sim$-connected model will be shown after the definition of the canonical model.

Definition 4.27 (Canonical Model). The canonical model M_Ω^Λ for a Λ-consistent set Ω of formulas from language $\mathcal{L}_{n,\bar{\sigma}}$ with extension $\mathcal{L}_{n,\bar{\sigma}}^+$ is the quadruple

$$\left\langle W_\Omega^\Lambda, (\sim_i^\Lambda)_{i \in I}, Dom^\Lambda, \mathcal{I}^\Lambda \right\rangle$$

where

1. W_Ω^Λ is the set of all Λ-MCSs with the $\forall$-property in $\mathcal{L}_{n,\bar{\sigma}}^+$ $\sim$-connected to some Λ-MCS extending Ω.

2. Each $\sim_i^\Lambda$ is the binary relation on W_Ω^Λ defined by $(w, w') \in \sim_i^\Lambda$ iff for every formula $K_i\varphi \in \mathcal{L}_{n,\bar{\sigma}}^+$, if $K_i\varphi \in w$, then $\varphi \in w'$.

3. $Dom^\Lambda = \bigcup_{\sigma \in \bar{\sigma}} Dom_\sigma^\Lambda$, where each $Dom_\sigma^\Lambda = \{[x] : x \in VAR_\sigma^+\}$ where $[x] = \{y : x \simeq y\}$ where $x \simeq y$ iff $(x = y) \in w$ for any $w \in W_\Omega^\Lambda$.

4. $\mathcal{I}^\Lambda$ is an interpretation such that

   a) Where P is a predicate symbol of arity $\sigma_1 \times \cdots \times \sigma_n$,

   $$\mathcal{I}^\Lambda(P, w) = \left\{ ([x_1], ..., [x_n]) \in Dom_{\sigma_1}^\Lambda \times \cdots \times Dom_{\sigma_n}^\Lambda : P(x_1, ..., x_n) \in w \right\}$$

   b) Where f is a function symbol of arity $\sigma_1 \times \cdots \times \sigma_n \longmapsto \sigma$,

   $$\mathcal{I}^\Lambda(f, w) = \left\{ ([x_1], ..., [x_n], [x]) \in Dom_{\sigma_1}^\Lambda \times \cdots \times Dom_{\sigma_n}^\Lambda \times Dom_\sigma^\Lambda : \right.$$
   $$\left. f(x_1, ..., x_n, x) \in w \right\}$$

c) Where c is a non-rigid constant, $\mathcal{I}^{\Lambda}(c, w) = [x] \in Dom^{\Lambda} : (x = c) \in w$.

d) Where c is a rigid constant, $\mathcal{I}^{\Lambda}(c) = [x] \in Dom^{\Lambda} : (x = c) \in w$, for any $w \in W_{\Omega}^{\Lambda}$. ◀

The choice of defining the domain of the canonical model as consisting of equivalence classes has been adopted from (Ebbinghaus et al., 1994). This choice is partly motivated partly by aesthetics (compare with (Hughes and Cresswell, 1996, p. 315-316)) and partly as it ensures that every term has a well-defined extension.

Definition 4.28 (Canonical Valuation). Define the *canonical valuation v^{Λ}* by $v^{\Lambda}(x) = [x]$. ◀

It will now be shown that all Λ-MCSs of a canonical model W_{Ω}^{Λ} share the same identity statements between variables.

Proposition 4.8 (Well-defined Domain). *For any $w, w' \in W_{\Omega}^{\Lambda}$ and any rigid terms $t, t' \in \mathcal{L}_{n,\bar{\sigma}}^{+}$, $(t = t') \in w$ iff $(t = t') \in w'$*

Proof. Let $w, w' \in W_{\Omega}^{\Lambda}$. Assume $w \sim_{i}^{\Lambda} w'$ and $(t = t') \in w$. Then by KI, $K_{i}(t = t') \in w$ and hence $(t = t') \in w'$ by the definition of $\sim_{i}^{\Lambda}$. Assume $w' \sim_{i}^{\Lambda} w$ and $(t = t') \in w$. If $(t = t') \notin w'$, then $(t \neq t') \in w'$ by maximality. By KNI, $K_{i}(t \neq t') \in w'$. Hence $(t \neq t') \in w$, which contradicts the consistency. If neither $w \sim_{i}^{\Lambda} w'$ nor $w' \sim_{i}^{\Lambda} w$ for any i, then w, w' are linked through other w''(s). As just argued, any two connected worlds share the same identity statements between rigid terms, and hence they all will. □

4.5.3 Key Lemmas: Existence and Truth

The following lemma is essential to the proof of the Truth Lemma below and has by far the most elaborate proof of any of the propositions in this text. The lemma guarantees that whenever a formula of the type $P_{i}\varphi$ is true at a state Γ in the canonical model, there will exist a suitable state w in the model related to Γ that satisfies φ. This is done in three steps. First, one constructs a suitable w relative to Γ by first constructing a consistent set w^{-} that contains the relevant formula φ. Secondly, it is then shown that w^{-} can be extended to a consistent set w^{+} that has the $\forall$-property. Finally, this set can be extended to a maximally consistent set w, which by construction is related to Γ.

Lemma 4.5 (Existence). *If $\Gamma \subseteq \mathcal{L}_{n,\bar{\sigma}}^{+}$ is a Λ-MCS with the $\forall$-property extending some formula set $\Omega \subseteq \mathcal{L}_{n,\bar{\sigma}}^{+}$ and $P_{i}\varphi \in \Gamma$, then there exists a $w \in W_{\Omega}^{\Lambda}$ such that $\Gamma \sim_{i}^{\Lambda} w$ and $\varphi \in w$.*

Proof. For some $i \in I$ assume that $P_{i}\varphi \in \Gamma$. A suitable w is constructed. Let w^{-} be $\{\varphi\} \cup \{\psi : K_{i}\psi \in \Gamma\}$.

First, w^- is Λ-consistent, as shown thus: assume w^- is Λ-inconsistent. As $\{\psi : K_i\psi \in \Gamma\}$ is Λ-consistent, then for some $\psi_1, ..., \psi_n \in \{\psi : K_i\psi \in \Gamma\}$,

$$\vdash_\Lambda (\psi_1 \wedge ... \wedge \psi_n) \to \neg\varphi$$

and by applying Gen, the obvious instance of K and MP it follows that that

$$\vdash_\Lambda K_i (\psi_1 \wedge ... \wedge \psi_n) \to K_i\neg\varphi$$

So by *K-distribution*

$$\vdash_\Lambda (K_i\psi_1 \wedge ... \wedge K_i\psi_n) \to K_i\neg\varphi$$

As Γ is assumed to be an Λ-MCS, $(K_i\psi_1 \wedge ... \wedge K_i\psi_n) \in \Gamma$ as $K_i\psi_1, ..., K_i\psi_n \in \Gamma$ (Prop. MCS). Thus, by MP, $K_i\neg\varphi \in \Gamma$, so, by Dual, $\neg P_i\varphi \in \Gamma$, but this violates the assumption that Γ is Λ-consistent. Hence, w^- is Λ-consistent.

Secondly, it is argued that from w^- a set w^+ can be constructed which has the $\forall$-property. To see this, define two enumerations. First, let there be given some enumeration of all formulas of the form $\forall y\psi$. Secondly, define a sequence of formulas $\varphi_0, \varphi_1, \varphi_2, ...$, where φ from above is φ_0. Let $\forall y\psi$ be the $n + 1$th formula in the first enumeration and z the first variable such that

$$\{\psi : K_i\psi \in \Gamma\} \cup \{\varphi_n \wedge (\psi(z/y) \to \forall y\psi)\} \qquad (*)$$

is Λ-consistent. Then define φ_{n+1} as $\varphi_n \wedge (\psi(z/y) \to \forall y\psi)$. Above it was shown that $w^- = \{\psi : K_i\psi \in \Gamma\} \cup \{\varphi_0\}$ is consistent. It is now shown that given $\{\psi : K_i\psi \in \Gamma\} \cup \{\varphi_n\}$ is consistent, there will always be a variable z satisfying $(*)$.[4]

Assume that there where no such z. Then for every variable z in $\mathcal{L}_{n,\overline{\sigma}}^+$, there would be some set of formulas $\{\chi_1, \chi_2, ..., \chi_n\} \subseteq \{\psi : K_i\psi \in \Gamma\}$ such that

$$\vdash_\Lambda (\chi_1 \wedge \chi_2 \wedge ... \wedge \chi_n) \to (\varphi_n \to (\psi(z/y) \wedge \neg\forall y\psi))$$

From *Derived Rule 1* it follows that

$$\vdash_\Lambda K_i (\chi_1 \wedge \chi_2 \wedge ... \wedge \chi_n) \to K_i (\varphi_n \to (\psi(z/y) \wedge \neg\forall y\psi))$$

and so, by *K-distribution*

$$\vdash_\Lambda (K_i\chi_1 \wedge K_i\chi_2 \wedge ... \wedge K_i\chi_n) \to K_i (\varphi_n \to (\psi(z/y) \wedge \neg\forall y\psi))$$

But as $\{\chi_1, \chi_2, ..., \chi_n\} \subseteq \{\psi : K_i\psi \in \Gamma\}$, the inclusion $\{K_i\chi_1, K_i\chi_2, ..., K_i\chi_n\} \subseteq \Gamma$ holds. As Γ is an Λ-MCS, it follows that $K_i (\varphi_n \to (\psi(z/x) \wedge \neg\forall y\psi)) \in \Gamma$ for all z.

Continuing, let z be a variable not occurring in ψ nor φ_n and consider the formula

$$\forall z K_i (\varphi_n \to \neg(\psi(z/y) \to \forall y\psi))$$

---

[4]Unfortunately, the saturation lemma cannot be used, as $\{\psi : K_i\psi \in \Gamma\}$ might contain all the variables of $\mathcal{L}_{n,\overline{\sigma}}^+$ in the first place.

As Γ has the $\forall$-property by assumption, there is some variable z' such that

$$K_i \left(\varphi_n \to \neg \left(\psi \left(z'/y\right) \to \forall y \psi\right)\right) \to \forall z K_i \left(\varphi_n \to \neg \left(\psi \left(z/y\right) \to \forall y \psi\right)\right)$$

belongs to Γ. But as $K_i \left(\varphi_n \to \left(\psi \left(z/x\right) \wedge \neg \forall y \psi\right)\right) \in \Gamma$ for all z, it also does so for z', so it follows that also $\forall z K_i \left(\varphi_n \to \neg \left(\psi \left(z/y\right) \to \forall y \psi\right)\right) \in \Gamma$.

Notice that

$$\forall z K_i \left(\varphi_n \to \neg \left(\psi \left(z/y\right) \to \forall y \psi\right)\right) \to K_i \forall z \left(\varphi_n \to \neg \left(\psi \left(z/y\right) \to \forall y \psi\right)\right)$$

is an instance of the Barcan Formula and is therefore also in Γ. Again, it follows that

$$K_i \forall z \left(\varphi_n \to \neg \left(\psi \left(z/y\right) \to \forall y \psi\right)\right) \in \Gamma.$$

From here, as z does not occur in φ_n, it follows by first-order reasoning that

$$K_i \left(\varphi_n \to \forall z \neg \left(\psi \left(z/y\right) \to \forall y \psi\right)\right) \in \Gamma. \tag{$**$}$$

Recall that $\exists y \left(\varphi \left(y/x\right) \to \forall x \varphi\right)$ for all formulas φ is a theorem of first-order logic if y is not free in $\forall x \varphi$,[5] from which it follows that

$$\vdash_\Lambda \exists z \left(\psi \left(z/y\right) \to \forall y \psi\right)$$

so $\exists z \left(\psi \left(z/y\right) \to \forall y \psi\right) \in \Gamma$. But from this and $(**)$ and the fact that Γ is Λ-consistent, $K_i \neg \varphi_n$ must be in Γ. But then $\{\psi : K_i \psi \in \Gamma\} \cup \{\varphi_n\}$ is Λ-inconsistent after all! Therefore conclude, that there always will be variable z making

$$\{\psi : K_i \psi \in \Gamma\} \cup \{\varphi_n \wedge \left(\psi \left(z/y\right) \to \forall y \psi\right)\} = \{\psi : K_i \psi \in \Gamma\} \cup \{\varphi_{n+1}\} \tag{$*$}$$

Λ-consistent.

Finally, $w^+ = \{\psi : K_i \psi \in \Gamma\} \cup \{\cup_{n \geq 0} \{\varphi_n\}\}$ is Λ-consistent since $\{\psi : K_i \psi \in \Gamma\} \cup \{\varphi_n\}$ for all n by the above is consistent and since $\vdash_\Lambda \varphi_m \to \varphi_n$ for $m \geq n$. Further, w^+ has the $\forall$-property by construction, and since w^+ is Λ-consistent it can be extended to a Λ-MCS w by Lindenbaum's Lemma. w also has the $\forall$-property by $\forall$-preservation. As $\{\psi : K_i \psi \in \Gamma\} \subseteq w$, by the definition of $\sim_i^\Lambda$, $\Gamma \sim_i^\Lambda w$. $\qquad\square$

Lemma 4.6 (Truth). *For every formula $\varphi \in \mathcal{L}_{n,\bar{\sigma}}$ and every $w \in W_\Omega^\Lambda$, $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \varphi$ iff $\varphi \in w$.*

Proof. The proof is by induction on the construction of φ. Let $w \in W_\Omega^\Lambda$.

Identity Let $t_1, t_2 \in TER_\sigma$. Then $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \left(t_1 = t_2\right)$ iff $[\![t_2]\!]_{v^\Lambda}^w = [\![t_2]\!]_{v^\Lambda}^w$ iff $[\![t_1]\!]_{v^\Lambda}^w, [\![t_2]\!]_{v^\Lambda}^w \in [x]$ and for some x iff $\left(t_1 = x\right), \left(t_2 = x\right) \in w$ iff $\left(t_1 = t_2\right) \in w$.

Atomic Let P be any predicate symbol of arity $\sigma_1 \times \cdots \times \sigma_n$ and let $t_k \in TER_{\sigma_k}$ for $k = 1, ..., n$. Then $M_\Omega^\Lambda, w \vDash_{v^\Lambda} P \left(t_1, ..., t_n\right)$ iff $\left([\![t_1]\!]_{v^\Lambda}^w, ..., [\![t_n]\!]_{v^\Lambda}^w\right) = \left([x_1], ..., [x_n]\right) \in \mathcal{I}^\Lambda \left(P, w\right)$ iff $P \left(x_1, ..., x_n\right) \in w$. But as $\left(t_1 = x_1\right), ..., \left(t_n = x_n\right) \in w$, $P \left(t_1, ..., t_n\right) \in w$.

---

[5](Hughes and Cresswell, 1996, p. 242)

Negation $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \neg\varphi$ iff not $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \varphi$ iff $\varphi \notin w$ iff $\neg\varphi \in w$

Conjunction $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \varphi \wedge \psi$ iff $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \varphi$ and $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \psi$ iff $\varphi \in w$ and $\psi \in w$ iff (Prop. MCS) $\varphi \wedge \psi \in w$.

Universal If Assume $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \forall x \varphi(x)$. Then for all x-variants $v^{\Lambda\prime}$ of v^Λ, $M_\Omega^\Lambda, w \vDash_{v^{\Lambda\prime}} \varphi(x)$. So, for all variables y, $M_\Omega^\Lambda, w \vDash_{v^{\Lambda\prime}} \varphi(y/x)$. By the induction hypothesis, $\varphi(y/x) \in w$ for all such y, especially for the y such that $(\varphi(y/x) \rightarrow \forall x \varphi) \in w$, which exist by the $\forall$-property. So $\forall x \varphi \in \varphi$.

Universal Only If By contraposition. Assume $\forall x \varphi(x) \notin w$. Then by maximality, $\neg \forall x \varphi(x) \in w$. Thus, by a contrapositive application of the $\forall$-property for some y, $\neg\varphi(y/x) \in w$. Hence, by the induction hypothesis, $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \neg\varphi(y/x)$. Thus, for the x-alternative to v^Λ where $v^{\Lambda\prime}(x) = v^\Lambda(y)$, $M_\Omega^\Lambda, w \vDash_{v^{\Lambda\prime}} \neg\varphi(x)$ so $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \exists x \neg\varphi(x)$ so $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \neg \forall x \varphi(x)$.

Modal If Assume that $K_i\varphi \in w$ and that $w \sim_i^\Lambda w'$ for arbitrary w'. Then by definition of $\sim_i^\Lambda$, $\varphi \in w'$. By the induction hypothesis, $M_\Omega^\Lambda, w' \vDash_{v^\Lambda} \varphi$ – this holds for all w' such that $w \sim_i^\Lambda w'$ as w' was arbitrary. Hence $M_\Omega^\Lambda, w \vDash_{v^\Lambda} K_i\varphi$.

Modal Only If Assume that $K_i\varphi \notin w$. Then $\neg K_i\varphi \in w$, so $P_i\neg\varphi \in w$. Hence, from the existence lemma, there is some $w' \in W_\Omega^\Lambda$ such that $\neg\varphi \in w'$ and $w \sim_i^\Lambda w'$. By the induction hypothesis, $M_\Omega^\Lambda, w' \vDash_{v^\Lambda} \neg\varphi$, and so $M_\Omega^\Lambda, w \vDash_{v^\Lambda} P_i\neg\varphi$ why $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \neg K_i\varphi$.

$\square$

4.5.4 Completeness Results

The canonical model for logic Λ was above defined relative to the set Ω of formulas to be shown satisfiable in a given instance.[6] This was done as the instance of the canonical model used is required to be connected in order to ensure the same identity statements holding between variables across all states. In order to reach the main result of this section, a last definition is required.

Above, the canonical model was defined relative to a Λ-consistent set Ω, so one last definition is required before stating the main result.

Definition 4.29 (**Canonical Class**)**.** For a normal $(n, \overline{\sigma})$ modal logic Λ define the class M^Λ of canonical models for Λ as the set of models M_Ω^Λ where Ω is any Λ-consistent set. ◀

Theorem 4.2 (**Canonical Class Theorem**)**.** *Any normal $(n, \overline{\sigma})$ modal logic Λ is (strongly) complete with respect to the class of canonical models for Λ.*

---

[6]This is opposed to the definition in (Blackburn et al., 2001, p. 197) where the canonical model for a propositional modal logic Λ includes *all* Λ-MCSs, resulting in *one* canonical model per logic.

Proof. By the proposition IFF above, the completeness of a normal $(n, \overline{\sigma})$ modal logic Λ with respect to a class of structures S reduces to providing, for each Λ-consistent set of formulas Ω, a structure $s \in \mathsf{S}$ such that s satisfies Ω. By Lindenbaum's Lemma and the Saturation Lemma, Ω can be extended to a Λ-MCS w with the $\forall$-property, and a model $M_\Omega^\Lambda \in \mathsf{M}^\Lambda$ can be defined such that $\Omega \subseteq w \in W_\Omega^\Lambda$. By the Truth Lemma above, $M_\Omega^\Lambda, w \vDash_{v^\Lambda} \Omega$. So, for any Λ-consistent set of formulas Ω, Ω is satisfiable on some structure $M_\Omega^\Lambda \in \mathsf{M}^\Lambda$, and thus Λ is complete with respect to the class of it's canonical models. $\qquad\square$

One easy application of the above theorem is to prove the completeness of $\mathbf{K}_{n,\overline{\sigma}}$ with respect to the class of all **n**-frames.

Theorem 4.3. *$\boldsymbol{K}_{n,\overline{\sigma}}$ is complete with respect to the class of all $(n, \overline{\sigma})$-frames.*

Proof. By IFF on page 61, it is sufficient to find for any $\mathbf{K}_{n,\overline{\sigma}}$-consistent set Ω a state w in a model M based on some $(n, \overline{\sigma})$-frame from $\mathsf{F}_{n,\overline{\sigma}}$, the class of all $(n, \overline{\sigma})$-frames, and some valuation v, such that $M, w \vDash_v \Omega$. Choose the model to be $M_\Omega^{\mathbf{K}_{n,\overline{\sigma}}}$, the state to be some $\mathbf{K}_{n,\overline{\sigma}}$-MCS Γ in $W_\Omega^{\mathbf{K}_{n,\overline{\sigma}}}$ extending Ω and the valuation to be $v^{\mathbf{K}_{n,\overline{\sigma}}}$. Then $M_\Omega^{\mathbf{K}_{n,\overline{\sigma}}}, \Gamma \vDash_{v^{\mathbf{K}_{n,\overline{\sigma}}}} \Omega$ by the Truth Lemma. Hence $\mathbf{K}_{n,\overline{\sigma}}$ is complete with respect to the class of all $(n, \overline{\sigma})$-frames by the Canonical Class Theorem. $\qquad\square$

Corollary 4.1. $\vdash_{\boldsymbol{K}_{n,\overline{\sigma}}} \varphi$ *if, and only if,* $\vDash_{\mathsf{F}_{(n,\overline{\sigma})}} \varphi$.

Proof. This follows immediately from Soundness and the previous theorem, taking Ω to be the empty set. $\qquad\square$

4.6 Adding Additional Axioms

In this section, the additional axioms T and 5 are introduced in order to define the logic $\mathbf{S5}_{n,\overline{\sigma}}$. This logic is then shown to be sound and complete with respect to the class $\mathsf{EQ}_{n,\overline{\sigma}}$ of $(n, \overline{\sigma})$-frames in with the accessibility relations are equivalence relations. As a special case, the completeness of the logic $\mathsf{QS5}$ with respect to the class $\mathcal{C}_{\mathsf{QEL}}$ of quantified epistemic models from chapter 3 follows.

Definition 4.30 (Axioms T and 5). Denote the following two axioms T and 5, respectively:

$$K_i\varphi \to \varphi \tag{T}$$
$$\neg K_i\varphi \to K_i\neg K_i\varphi \tag{5}$$

$\blacktriangleleft$

Notice that $P_i\psi \to K_iP_i\psi$ is an instance of 5 using the definition of P_i and $\varphi := \neg\psi$. This is instance is used in the proof for $\mathbf{S5}_{n,\overline{\sigma}}$ completeness below.

Definition 4.31 (Logic $S5_{n,\overline{\sigma}}$). The logic $\mathbf{S5}_{n,\overline{\sigma}}$ based on language $\mathcal{L}_{n,\overline{\sigma}}$ is the smallest set of $\mathcal{L}_{n,\overline{\sigma}}$ formulas obtained by adding to the $\mathbf{K}_{n,\overline{\sigma}}$ the axioms T and 5, and which is closed under the $\mathbf{K}_{n,\overline{\sigma}}$ inference rules. ◀

Definition 4.32 (Class $EQ_{n,\overline{\sigma}}$). Let $EQ_{n,\overline{\sigma}}$ denote the class consisting of all $(n,\overline{\sigma})$-frames in which all accessibility relations are equivalence relations. ◀

Lemma 4.7. *Where M_Ω^Λ is a canonical model as defined in Definition 4.27, w, w' are states in M_Ω^Λ and ψ any $\mathcal{L}_{n,\overline{\sigma}}^+$ formula:*

$$w \sim_i^\Lambda w' \text{ iff } (\psi \in w' \text{ implies } P_i\psi \in w)$$

Proof. Assume that $w \sim_i^\Lambda w'$. By Definition 4.27, $w \sim_i^\Lambda w'$ iff ($K_i\varphi \in w$ implies $\varphi \in w'$), for all possible formulas $K_i\varphi \in \mathcal{L}_{n,\overline{\sigma}}^+$. This is equivalent to the requirement that ($\neg\varphi \in w'$ implies $\neg K_i\varphi \in w$). Rewritten using the definition of P_i, this amounts to ($\neg\varphi \in w'$ implies $P_i\neg\varphi \in w$). Setting $\varphi := \neg\psi$ proves the lemma. □

Theorem 4.4 ($S5_{n,\overline{\sigma}}$ Soundness). *The logic $\mathbf{S5}_{n,\overline{\sigma}}$ is sound with respect to the class $EQ_{n,\overline{\sigma}}$.*

Proof. It must be shown that all axioms of $\mathbf{S5}_{n,\overline{\sigma}}$ are valid in $EQ_{n,\overline{\sigma}}$, and that the $\mathbf{S5}_{n,\overline{\sigma}}$ inference rules preserve validity in the same class.

Lemma 4.1 shows that all $\mathbf{K}_{n,\overline{\sigma}}$ axioms are valid in all frames, hence also in all $EQ_{n,\overline{\sigma}}$ frames. Further, Lemma 4.2 shows that all $\mathbf{K}_{n,\overline{\sigma}}$ inference rules preserve truth on all frames, hence also on all $EQ_{n,\overline{\sigma}}$ frames. Hence, to show that $\mathbf{S5}_{n,\overline{\sigma}}$ is sound with respect to $EQ_{n,\overline{\sigma}}$, the only thing required is to show that the axioms T and 5 are valid in this class.

T. Assume $\mathcal{F} = \langle W, (\sim_i)_{i \in I}, Dom \rangle$ is a $EQ_{n,\overline{\sigma}}$ frame, and that $M, w \models_v K_i\varphi$ for any model, world, valuation, i and φ. Then, for all w' such that $w \sim_i w'$, $M, w' \models_v \varphi$. As $\sim_i$ is an equivalence relation, it is also reflexive. Therefore, $w \sim_i w$. Hence $M, w \models_v \varphi$. From this it follows that $M, w \models_v K_i\varphi \to \varphi$. As M, w, v and i were arbitrary, T is shown to be valid in $EQ_{n,\overline{\sigma}}$.

5. To show 5 is valid in $EQ_{n,\overline{\sigma}}$, assume that M is a model based on a $EQ_{n,\overline{\sigma}}$ frame. Then $\sim_i$ is euclidean, for all $i \in I$. Assume further that $M, w \models_v \neg K_i\varphi$, for some w, v and i. By the existence lemma and the definition of the P_i-operator, there is some w' such that $w \sim_i w'$ and $M, w' \models_v \neg\varphi$. Assume for a contradiction that $M, w \models_v \neg K_i \neg K_i\varphi$. Then there is some w'' such that $w \sim_i w''$ and $M, w'' \models_v K_i\varphi$, as $\neg K_i \neg K_i\varphi$ and $P_i K_i\varphi$ are equivalent. As $w \sim_i w''$ and $w \sim_i w'$ it follows by the assumption of euclideaness that $w'' \sim_i w'$. As $M, w'' \models_v K_i\varphi$, it hence follows that $M, w' \models_v \varphi$, and a contradiction is reached. It is therefore concluded that $M, w \models_v K_i \neg K_i\varphi$, to the effect that $M, w \models_v \neg K_i\varphi \to K_i \neg K_i\varphi$. As M was arbitrary, 5 is valid in the class $EQ_{n,\overline{\sigma}}$.

It is concluded that T and 5 are both valid in the class $EQ_{n,\overline{\sigma}}$, and the results is therefore proven. □

To avoid cumbersome notation, for the following completeness proof, a notational shorthand will be used: the canonical relation $\sim_i^{\mathbf{S5}_{n,\overline{\sigma}}}$ will be written $\sim_i^*$.

Theorem 4.5 ($S5_{n,\bar\sigma}$Completeness). *The logic $S5_{n,\bar\sigma}$ is complete with respect to the class $EQ_{n,\bar\sigma}$.*

Proof. By the Canonical Model Theorem, the only thing required is to show that the accessibility relations in the class of canonical models for $\mathbf{S5}_{n,\bar\sigma}$ are equivalence relations. This is done in two steps: first they are shown to be reflexive, using the axiom T. Second, they are shown to be euclidean, using axiom 5.

Reflexive. Let Ω be any $\mathbf{S5}_{n,\bar\sigma}$-consistent set of formulas, and let $M_\Omega^{\mathbf{S5}_{n,\bar\sigma}}$ the canonical model for this set, as specified in Definition 4.27. Let w be a state of this model, and assume that $K_i\varphi \in w$. As w is a $\mathbf{S5}_{n,\bar\sigma}$-MCS, it contains T: $K_i\varphi \to \varphi$. By the consistency requirement, $\varphi \in w$. As φ was arbitrary, it can be concluded by the definition of $\sim_i^*$ (i.e. $\sim_i^{*\mathbf{S5}_{n,\bar\sigma}}$) that $w \sim_i^* w$. As φ, w and i were chosen arbitrarily, it can be concluded that all accessibility relations in $M_\Omega^{\mathbf{S5}_{n,\bar\sigma}}$ are reflexive. As Ω was arbitrary, the class of canonical models for $\mathbf{S5}_{n,\bar\sigma}$ is reflexive.

Euclidean. Let Ω and $M_\Omega^{\mathbf{S5}_{n,\bar\sigma}}$ be as above, and assume that w_0, w_1 and w_2 are states in $M_\Omega^{\mathbf{S5}_{n,\bar\sigma}}$ such that $w_0 \sim_i^* w_1$ and $w_0 \sim_i^* w_2$, for some $i \in I$. It will be shown that $w_2 \sim_i^* w_1$.

Notice that $P_i\psi \to K_iP_i\psi$ is an instance of 5 using the definition of P_i and $\varphi := \neg\psi$. Now, let φ be some formula such that $\varphi \in w_1$. By Lemma 4.7 this implies that $P_i\varphi \in w_0$. As w_0 is a $\mathbf{S5}_{n,\bar\sigma}$-consistent set, $P_i \to K_iP_i\varphi \in w_0$. Utilizing modus ponens, this implies that $K_iP_i\varphi \in w_0$. By the definition of $\sim_i^*$ and the assumption that $w_0 \sim_i^* w_2$, it follows that $P_i\varphi \in w_2$. Hence, by Lemma 4.7, $w_2 \sim_i^* w_1$.

As Ω and $M_\Omega^{\mathbf{S5}_{n,\bar\sigma}}$ in both cases were chosen arbitrarily, it is concluded that the accessibility relations of any $M_\Omega^{\mathbf{S5}_{n,\bar\sigma}}$ model is reflexive and euclidean. Hence the class $\mathsf{M}^{\mathbf{S5}_{n,\bar\sigma}}$ of $\mathbf{S5}_{n,\bar\sigma}$ canonical models is included in the class of $EQ_{n,\bar\sigma}$, and the result is shown. $\qquad\square$

Corollary 4.2 (QS5 Completeness). *The logic QS5 is sound and complete with respect to the class $\mathcal{C}_{QEL}$.*

Proof. If based on the same language $\mathcal{L}_{n,\bar\sigma}$, the logic QS5 is equal to the logic $\mathbf{S5}_{n,\bar\sigma}$ with $\bar\sigma = \{\sigma\}$, where σ is a non-rigid sort. By Theorem 4.4 and Theorem 4.5, it is therefore sound and complete with respect to the class $EQ_{n,\bar\sigma}$. Soundness with respect to $\mathcal{C}_{QEL}$ follows as each $\mathcal{C}_{QEL}$ model is based on an $EQ_{n,\bar\sigma}$ frame. Completeness follows as as each $\mathcal{C}_{QEL}$ model is based on an $EQ_{n,\bar\sigma}$ frame, why $\Gamma \models_{EQ_{n,\bar\sigma}} \varphi$ implies $\Gamma \models_{\mathcal{C}_{QEL}} \varphi$ for all sets of formulas Γ and formulas φ of $\mathcal{L}_{n,\bar\sigma}$. Hence, for same Γ and φ, $\Gamma \models_{\mathcal{C}_{QEL}} \varphi$ implies $\Gamma \vdash_{\mathbf{S5}_{n,\bar\sigma}} \varphi$. $\qquad\square$

5 Modeling Lexical Competence

In order to construct a formal theory of semantic competence, a formal counterpart to Marconi's structure of lexical competence (SLC) as presented in section 2.2.3 is constructed, validated and explored. This is the topic of this chapter, and is done as follows. In section 5.1, some simplifying assumptions regarding the SLC are made, and an overview of the model is presented. In section 5.2, the QEL framework is extended to accommodate the addition of a word lexicon. Once the model is constructed, the topic turns to two issues of validation. One is to ensure proper correspondence between the elements of the conceptual and formal theories, and the second is to show property conservation. In section 5.3, the first issue is undertaken. The three ontologies of Marconi's conceptual theory are identified in the formal model, both semantically and syntactically, and various properties are discussed in relation to QEL. The conclusion is that proper formal counterparts to the ontologies are present. The second validation issue is discussed in section 5.4, where the three competence types are identified. It is shown that the dissociation of these competence types are conserved in the model. The model therefore preserves all essential elements of Marconi's theory, but in a proper formal framework. In section 5.5, this formal framework is investigated, focusing on topics not discussed by Marconi.

5.1 Overview and Simplifying Assumptions

To model SLC, a stronger framework than quantified epistemic logic is required. The reason for this is the dissociation of the semantic lexicon and word lexicas, as expounded in section 2.2.3. In particular, if the word lexicas were removed from SLC, so this represented only a conceptual model of agents cognition and the world, the QEL framework would be sufficient, and could even be seen to give a more detailed analysis of this structure in terms of knowledge. Adding the word lexicas requires an additional aspect not included in QEL.

The strategy for modeling SLC is as follows. First, attention is restricted to objects, semantic lexicas consistent of objects and word lexicas including only unambiguous proper names. In the same vein, the entries in the semantic lexicon will consist only of individual concepts.

Regarding real-world objects, these will correspond on the semantic side to objects in the domain of quantification. On the syntactic side, objects are thus represented as the ordinary constants and variables. The entries in a semantic lexicon will be represented by the agents individual concepts. A formal definition of the individ-

ual concepts is given on the semantic level, and it is shown how this can be expressed syntactically. The semantic counterpart to individual concepts are termed *object indistinguishability classes.* As QEL cannot accommodate the word lexica, the framework will be augmented with an additional sort of constants and variables functioning as a word lexicon. To match these semantically, an additional domain is introduced. This domain is interpreted as a catalog of *proper names.*

As mentioned in section 2.2.3, two word lexica are included in SLC, but more could readily be incorporated. Hence, the number of word lexica included is not a critical aspect of the structure. To simplify the modeling process and the model itself, the focus will be on a structure restricted to only containing one word lexicon. It will further be assumed that the agents are able to distinguish between different word types, i.e. that they are *syntactically competent* with respect to the words in the lexicon. Hence, whenever an agent is presented with two name tokens of the same type of name, the agent knows that these are tokens of the same name type. Identity between name tokens are interpreted stating that the tokens belong to the same type. Figure 5.1.1 illustrates the simplified SLC[1] and the formal apparatus corresponding to each module, which will be defined below.[2]

The content of the single word lexicon will be restricted to only consist of proper names. As will be expounded below, the words in the lexicon will have a formal first-order counterpart. In the case of proper names, this counterpart will be first-order constants. These word terms will be assigned Millian meaning by a mapping to the set of object terms. If verbs were included in the word lexicon, these would have predicates as their formal counterpart, and the mapping assigning meaning would have to be to the set of relation symbols. On the semantic side, this would require a mapping to the power set of the object domain. In order to express identity in meaning between verbs, second-order identity would hence be required. Thus, the word lexicon will not include a word for the natural language term equivalent to identity, i.e. "is". To stay within a first-order framework only proper names are included in the word lexicon. This does limit the model quite severely, in particular regarding aspects of inferential competence, but will still result in a theory capable of providing interesting insights. In particular, in order to analyze Frege's puzzles, this is sufficient.

The structure of this chapter is as follows: in section 5.2, the QEL framework will be augmented with an additional sort of constants and variables to represent the word lexicon. Additionally, a meaning function is defined in order to assign meanings to the proper names of the word lexicon. In section 5.3, the augmented structure will be compared to the SLC. Object indistinguishability classes are defined and discussed. In section 5.4, the competence types described by Marconi are identified in the formal framework and it is shown that they a dissociated. Finally, in section 5.5, selected validities of the augmented framework are reviewed and discussed. As cer-

---

[1]In the following, when talking about the structure of lexical competence, reference will be to this simplified version.

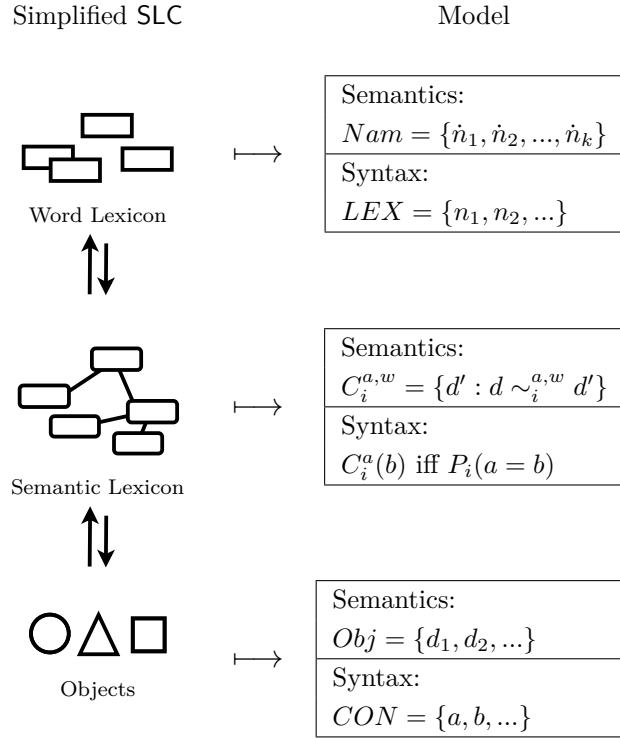[2]The figure is slightly misleading, but illustrative.

Figure 5.1.1: Simplified SLC and to be defined formal model. The word lexicon is modeled by adding to QEL a domain of *names* to the semantics and an additional sort to the syntax. The semantic lexicon is represented by *indistinguishability classes* of objects, and a modal *concept predicate* can be defined. Objects are modeled as in QEL.

tain implicational relationships exist between conjunctions of competence types, a hierarchy of competence levels is discussed. The reader not interested in this aspect may choose to skip this discussion as it is not used later, but included *only* for it's own sake.

5.2 Augmenting Quantified Epistemic Logic

In the following, the QEL framework will be extended. First, the language of QEL will be extended to a language $\mathcal{L}_{2QEL}$. Secondly, the class of QEL models is extended to construct a class of models with two-sorted domains. The augmented framework is denoted 2QEL.

In order to model the word lexicon as separate from the semantic lexicon and real-world objects, a new sort of terms is introduced. As the exposition is limited to only one lexicon, only one additional sort will be introduced. The logic will therefore be *two-sorted*. The first sort, denoted σ_{OBJ}, is that used in QEL, and denotes objects. σ_{OBJ} is the sort consisting of CON and VAR from QEL. The newly added sort σ_{LEX}

has constants LEX and variables VAR_{LEX}.[3] The details are presented below.

In order to add further word lexica, more rigid sorts should be added to the syntax, and the semantics and axiom system should be extended accordingly.

In this section, the additional sort will be introduced in the syntax, and a matching semantics will be given. As the formal interpretation of the constants of the second sort will be different from that of QEL, axioms are shortly mentioned and discussed, and completeness is shown.

5.2.1 Syntax and Semantics

In order to model the SLC, a proper language for doing so will be defined. This is done by extending the language of QEL with an additional sort. This means making a move to two-sorted epistemic logic. The sort used in chapter 3 will be denoted σ_{OBJ} – the object sort. This has constants CON and variables VAR. The set of object terms is denote TER_{OBJ}, which includes both CON and VAR.

The newly added sort is denoted σ_{LEX}, as this represents *lex*ical items. The set of σ_{LEX} constants is

$$LEX = \{n_1, n_2, ...\}$$

and the set of σ_{LEX} variables is

$$VAR_{LEX} = \{\dot{x}_1, \dot{x}_2, ...\}.$$

The set of lexical terms is denoted TER_{LEX}, and $LEX \cup VAR_{LEX} \subseteq TER_{LEX}$. The lexical terms is interpreted as *name tokens* (see below).

As mentioned in section 2.2.3, no reasoning occurs *in* the word lexicon, and no relations between the lexical items exist. In order to capture this feature, the modeling language will be restricted to include only predicates and relations for object terms. The only relation included for lexical terms is identity, as this is required for modeling *syntactical competence*, as explained below. This restriction results in the following definition of the modeling language:

Definition 5.1 (Modeling Language $\mathcal{L}_{2QEL}$). The modeling language $\mathcal{L}_{2QEL}$ consists of

1. Two sorts, σ_{OBJ} and σ_{LEX}, with assigned constants and variables as above.

2. A set REL of relation symbols. Each has an arity $n \in \mathbb{N}$ and takes n arguments from TER_{OBJ}.

3. The identity symbol $=$.

4. A function symbol μ of sort $\sigma_{LEX} \longmapsto \sigma_{OBJ}$.

5. Logical connectives, operators and quantifiers as in $\mathcal{L}_{QEL}$.   ◀

---

[3]In the terminology of the previous chapter, the logic used is $\mathsf{QS5}_{n,\overline{\sigma}}$, where $\overline{\sigma} = \{\sigma_{OBJ}, \sigma_{LEX}\}$ where σ_{LEX} is a rigid sort.

The inclusion of the the function symbol μ has so far gone unmentioned. The function is interpreted as a *meaning assigning function*, used to model Millian meaning (see below). The function assigns to each lexical term t an object term $\mu(t)$. Hence, where $t \in TER_{LEX}$ it is required that $\mu(t) \in TER_{OBJ}$.

Using Definition 5.1, the well-formed formulas of $\mathcal{L}_{\mathsf{2QEL}}$ may now be defined as follows:

Definition 5.2 ($\mathcal{L}_{\mathsf{2QEL}}$ **Well-Formed Formulas**)**.** Define the set of $\mathcal{L}_{\mathsf{2QEL}}$ *well-formed formulas* by

1. Where $t_1, t_2 \in TER_{LEX} \cup TER_{OBJ}$, the expression $(t_1 = t_2)$ is an *atomic* $\mathcal{L}_{\mathsf{2QEL}}$ *well-formed formula*.

2. Where R is a n-ary relation symbol and $t_1, ..., t_n \in TER_{OBJ}$, $R(t_1, ..., t_n)$ is an *atomic* $\mathcal{L}_{\mathsf{2QEL}}$ *well-formed formula*.

3. All atomic $\mathcal{L}_{\mathsf{2QEL}}$ well-formed formula are $\mathcal{L}_{\mathsf{2QEL}}$ well-formed formulas.

4. Where φ and ψ are $\mathcal{L}_{\mathsf{2QEL}}$ well-formed formula, $x \in VAR \cup VAR_{LEX}$ and $i \in I$, the following are $\mathcal{L}_{\mathsf{2QEL}}$ well-formed formulas

$$\neg\varphi | \varphi \wedge \psi | \forall x \varphi | K_i \varphi.$$

◀

Turning to the semantics, the structure of the $\mathcal{C}_{\mathsf{QEL}}$ models defined in section 3.2.2 must be augmented with an additional domain, consisting of *name types*. This domain is denoted Nam. The models will also contain an *object domain* with the same interpretation as the domain in QEL. This domain is denoted Obj. To simplify the truth clauses for quantification, the union of the two is referred to simply as the *domain*, denoted Dom.

Definition 5.3 (2QEL **Domains**)**.** The 2QEL domains consists of two disjunct sets

1. The *name domain* is a non-empty finite set, $Nam = \{\dot{n}_1, \dot{n}_2, ..., \dot{n}_k\}$.

2. The *object domain* is a non-empty countable set, $Obj = \{d_1, d_2, ...\}$.

and their union, the *domain*, $Dom = Nam \cup Obj$. ◀

The name domain is interpreted as consisting of name *types*, in contrast to the set of lexical terms, which are interpreted as name *tokens*. The set Nam can therefore be seen as the agents' dictionary: it contains one entry for every type of word available. The name domain is further assumed finite so agents with bounded resources are able to learn all the entries, see e.g. (Davidson, 1984).[4] Recall that in QEL models,

---

[4]Note: in the multi-agent case, the added domain will be common knowledge among all agents, an unrealistic default assumption. Hence the assumptions of a finite size may seem irrelevant as the domain is known due to the constant domain assumption, but as it is a common assumption in the philosophy of language, it will be implemented here, too.

the elements of the domain is thought of by the agents as *objects*, not names of such. Adding the name domain introduces *names* for such objects.

In order to assign values to both object and lexical terms, the interpretation function of QEL must be extended. The interpretation $\mathcal{I}$ must assign object constants values in the object domain, and lexical constants values in the name domain. Finally, the interpretation must assign a function on the domain to the meaning function symbol.

Definition 5.4 (2QEL **Interpretation**)**.** Define the 2QEL *interpretation* to be a map $\mathcal{I}$ where

1. $\mathcal{I}$ assigns to each n-ary relation symbol and world a n-ary relation on Obj, i.e.
$$\mathcal{I} : REL_n \times W \longrightarrow \mathcal{P}(Obj^n)$$

2. $\mathcal{I}$ assigns to each object constant and world an element in the object domain, i.e.
$$\mathcal{I} : CON \times W \longrightarrow Obj$$

3. $\mathcal{I}$ assigns surjectively to each name constant an element in the name domain, i.e.
$$\mathcal{I} : LEX \longrightarrow Nam$$

4. $\mathcal{I}$ assigns to the function symbol μ a function from Nam to Obj, relative to each world, i.e.
$$\mathcal{I} : \{\mu\} \times W \longrightarrow Obj^{Nam}$$

◀

The first two clauses in this definition are equivalent to the QEL case, but the two last deserve comment. Regarding (3), notice then that the assignment of names to lexical items is *not world relative.* This requirement is to make sure that the agents are *syntactically competent* with respect to their vocabulary, i.e. that they are able to distinguish names from one another. The requirement is warranted with the interpretation of the constants of LEX as *tokens* of names and the elements of Nam as *types*, and syntactical competence with a given name understood as the knowledge of the identity of two tokens of the same name. For then the semantics validate that whenever $n_1 = n_2$ it follows that $K_i (n_1 = n_2)$. Such identity statements between names *do not* convey any information regarding the meaning of the names. Rather, they express identity of the two *signs.* Hence, the identity '*London = London*' is true, where as the identity '*London = Londres*' is false – as the two first occurrences of '*London*' are two tokens of the same type, whereas the '*London*' and '*Londres*' are occurrences of two different name types, albeit with the same meaning. That $\mathcal{I}$ is assumed to be a surjective function from LEX to Nam is assumed so each name type has at least one name token.

In relation to (4), two important assumptions are explicitly made regarding the meaning assigning function, μ. The first is that *meaning is assigned by a function*, and the second is that *the assigned meaning is world-relative.* Both require comment.

Meaning assigned by a function In order to investigate the theory of Millian meaning, this theory must be embedded in the formal framework. This is simple due to the simplicity of the Millian theory: all it takes is for each name to be assigned a referent. As mentioned shortly in section 2.1.1, reference is typically assigned by some *theory of reference*. Such a theory has the role of clarifying the nature of the *reference relation*, being the relation by which names get their reference. The theory of reference utilized can for present purposes be "black boxed" as the present purpose does not involve evaluating the internal structure of theories of reference. The only assumption made regarding the reference relation is that it is in fact a *function*. This assumption is made since it is easier to work with *unambiguous* proper names. Assuming that referents are assigned by a function ensures exactly this: that each name is assigned *exactly* one referent, i.e. one meaning. Particularly, it is *not* assumed that μ is injective, as it is essential to be able to have co-referring terms.[5]

On the syntactic level, this is accomplished by adding the unary meaning function μ of arity $\sigma_{TER} \longmapsto \sigma_{OBJ}$. That is, μ assigns to each lexical term a meaning from the object terms. From the viewpoints of the agents, μ assigns an object to each name. On the semantic level, μ is assigned a function from *Nam* to *Obj*, relative to each world. This assumption is discussed below.

Adding meaning in this way is contrary to the viewpoints of Marconi. He explicitly states that "in [his] picture, meanings are nowhere to be found" (p. 81). Marconi does not include meanings as a matter of philosophical conviction, but his theory of lexical competence is compatible with an inclusion of such.

Meaning assigned world-relatively On the semantic level, the reference map is defined *world relatively*, as illustrated in figure Figure 5.2.1. This means that the value $\mu(n)$ for $n \in LEX$ can change from world to world. Hence, *names are assigned values relative to epistemic alternatives*.

The primary reason for assigning names relative to epistemic alternatives is that agents should not know the meaning of terms by default. In order to model that agents can be uncertain regarding the meaning of a term, the meaning must be able to change across the epistemic alternatives. If the meaning function was constant across all worlds, the agents would by default be able to identify the referent of every name. As this ability is a type of competence, it should be possible that the agents lack this knowledge. Hence, the meaning is assigned relative to epistemic alternatives.

That meaning is assigned world-relatively may seem to clinch with the common assumption in the philosophy of language after (Kripke, 1980) that names refer *rigidly*. That a name refers rigidly, or is a *rigid designator*, means that the name refers to the same object in *all metaphysically possible worlds*, cf. (Lycan, 2006). However, the world-relative meaning assignment does not clinch with this assump-

---

[5]It could be assumed that μ was a partial function, in order to allow for empty names. In order to not complicate matters, this is not done here.
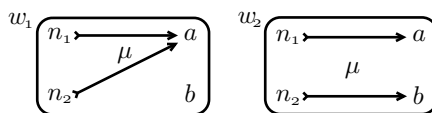
Figure 5.2.1: The meaning function μ is defined world relatively, i.e. the meaning of a name may shift across epistemic alternatives. In the portrayed model, the meaning of n_1 is constant, but n_2 denotes a in w_1 and b in w_2.

tion as the worlds over which the meaning varies are *epistemic alternatives*. These epistemic alternatives may be *metaphysically impossible*, without the agents knowing them to be so. The epistemic alternatives can deviate from the actual world in any *logically possible way*[6]. Only the actual world is assumed to be metaphysically possible. Across this singleton set, the function is always constant (trivially), and hence the names are rigid designators.

To be able to more succinctly refer to the semantic equivalent of $\mu(n)$, when $(\dot{n}, d) \in \mathcal{I}(\mu, w)$ and $\mathcal{I}(n) = \dot{n}$, this will be written $\mathcal{I}(\mu(n), w) = d$. That is, $\mathcal{I}(\mu(n), w) = d$ states that d is the referent of $\dot{n}$ at w.

Returning to the definitions of 2QEL semantics, the notion of a 2QEL model may now be defined. With the primary extensions given by the definitions above, the 2QEL models may now be defined as follows:

Definition 5.5 (2QEL **models**)**.** A 2QEL *model* M is a quintuple

$$M = \langle W, (\sim_i)_{i \in I}, Dom, \mathcal{I} \rangle$$

where W and $(\sim_i)_{i \in I}$ are as in the QEL case, $Dom = Obj \cup Nam$ is a 2QEL domain as defined in Definition 5.3, and $\mathcal{I}$ is a 2QEL interpretation as defined in Definition 5.4.

Where M is a 2QEL model $\langle W, (\sim_i)_{i \in I}, Dom, \mathcal{I} \rangle$ and $w \in W$, the pair (M, w) is a *pointed* 2QEL *model.*

The class of 2QEL models is denoted $\mathcal{C}_{2QEL}$. ◀

In order to assign truth-conditions to the formulas of $\mathcal{L}_{2QEL}$, the variables needs to be assigned extensions. Since and additional sort of variables has been added, the definition of the valuation must be altered to take these into account.

Definition 5.6 (2QEL *Valuation*)**.** A 2QEL *valuation* is a map

$$v : VAR \longrightarrow Obj$$
$$v : VAR_{LEX} \longrightarrow Nam$$

◀

The valuation assigns extensions to the variables of $\mathcal{L}_{2QEL}$. Specifically, it ensures that the object variables of VAR are assigned values in the object domain, Obj, and

---

[6]Based on the present axiom system.

that the lexical variables in VAR_{LEX} are assigned values in the name domain, Nam. The definition of an x-variant of a valuation v is identical to the QEL case.

The definitions of QEL truth-conditions can be altered to include formulas with lexical terms. The truth-conditions are to a high degree identical to the QEL truth-conditions. In fact, the only aspect that needs altering is the truth-conditions for atomic formulas. These can be defined as follows:

Definition 5.7 (**Truth-Conditions for** 2QEL **atomic formulas**)**.** Where (M, w) is a pointed 2QEL model with interpretation $\mathcal{I}$ and v is a 2QEL valuation, the satisfaction relation $\models$ for atomic formulas is given by the following clauses:

$$
\begin{aligned}
M, w \models_v P(t_1, t_2, ..., t_n) \quad &\text{iff} \quad (d_1, d_2, ..., d_n) \in \mathcal{I}(P, w) \\
M, w \models_v (t_1 = t_2) \quad &\text{iff} \quad d_1 = d_2 \\
&\text{where} \quad d_i = \begin{cases} v(t_i) & \text{if } t_i \in VAR \\ v(t_i) & \text{if } t_i \in VAR_{LEX} \\ \mathcal{I}(w, t_i) & \text{if } t_i \in CON \\ \mathcal{I}(t_i) & \text{if } t_i \in LEX \end{cases} \\
&\text{for} \quad i \in \{1, 2, ...\}
\end{aligned}
$$

◄

The truth-conditions for the remaining formulas are as stated in section 3.2.2, with all modifications mentioned here in effect. In particular, recall that in the two-sorted version, $Dom = Nam \cup Obj$.

Notice that in the given syntax and semantics of 2QEL, identity can be stipulated between a name and an object. Such identities will always be false, as the two sets are disjoint at the semantic level, and hence items from LEX and CON will never be interpreted as the same object in Dom.

This concludes the model-theoretic definitions of the 2QEL framework. This framework will be used as the foundation for a model-theoretic modeling of Marconi's SLC. Still, in order to provide a *logical* theory, a logic based on the 2QEL models must be defined.

Definition 5.8 (**The Logic** $\mathsf{L_{2QEL}}$)**.** Define the logic $\mathsf{L_{2QEL}}$ by

$$
\mathsf{L_{2QEL}} = \{\varphi \in \mathcal{L}_{2QEL} : M \models \varphi \text{ for all } M \in \mathcal{C}_{2QEL}\}
$$

◄

The definition states that the logic $\mathsf{L_{2QEL}}$ consists *exactly* of the formulas valid in all 2QEL models. This is a logic defined from the semantic viewpoint. From this definition, a formula may be determined as being in the logic by showing that it is a $\mathcal{C}_{2QEL}$ validity. If a $\mathcal{C}_{2QEL}$ model exists in which the negation of a formula φ can be satisfied, it can then be concluded that φ is *not* a theorem of $\mathsf{L_{2QEL}}$.

In order to provide a set of axioms and inference rules from which exactly the theorems of $\mathsf{L_{2QEL}}$ can be deduced, the following section introduces a selection of the appropriate axioms. For the full list, the reader is referred to the previous chapter. The axioms presented below correspond to important features of the 2QEL models.

5.2.2 Axioms

With respect to providing a complete axiomatization for $\mathsf{L_{2QEL}}$, the main difference between QEL and $\mathsf{2QEL}$ is that the latter includes a set of terms which are rigid by definition, namely the set TER_{LEX}. All QEL axioms are valid in $\mathsf{2QEL}$ models, but the axioms restricted to variables in the earlier case are now valid for variables $\dot{x} \in VAR_{LEX}$ and terms $t, t' \in TER_{LEX}$. This means that the following, stronger versions[7] of the QEL axioms UI, PS and KNI must be included in the axiomatic base for $\mathsf{L_{2QEL}}$:

$$
\begin{array}{ll}
\text{UI}^* & \forall \dot{x}\varphi \to \varphi\left(t/\dot{x}\right) \\
\text{PS}^* & (t = t') \to (\varphi\left(t\right) \leftrightarrow \varphi\left(t'\right)) \\
\text{KNI}^* & (t \neq t') \to K_i\left(t \neq t'\right)
\end{array}
$$

As the set TER_{LEX} consists of rigid terms, the system of identity used is *incontingent*. This implies that, for $t \in TER_{LEX}$, truth is preserved by the classic version of *Existential Generalization*:

$$
\frac{\varphi\left(t\right)}{\exists \dot{x}\varphi\left(\dot{x}\right)}
$$

This rule of inference is redundant by the addition of UI*, but nicely illustrates differences between the non-rigid object constants and the rigid lexical constants. For the former, the rule fails to preserve truth, whereas for the latter, the rule does preserve truth. Where $\varphi(t)$ is the formula $K_i(t = t_1)$ the rule can be used to illustrate an important difference between the two identity systems. Where t is an object term, it does not preserve truth to conclude that $\exists x K_i \varphi(x)$ – i.e. it cannot be concluded that i can identify t_1, as was discussed in section 3.2.3. On the other hand, where t is a lexical constant, it can be concluded that $\exists \dot{x} K_i(\dot{x} = t_1)$ – i.e. i can identify *the name* t_1. The truth of the latter is a feature of syntactical competence: where i knows that two name tokens are of the same type, the specific name can be identified.

   Another feature of syntactical competence is embodied in the derivability of the formula *Knowledge of Identicals*

$$
\text{KI}^* \quad (t = t') \to K_i(t = t')
$$

KI* states that for all lexical terms $t, t' \in TER_{LEX}$, if such two are tokens of the same name, then any agent knows this by default. In conjunction with Existential Generalization this implies that, by default, all agents always know that two name tokens of the same type are of the same type, and they are able to identify this type. Hence, agents can distinguish and identify all names.

---

[7]The following axioms are inspired by inspired by (Hughes and Cresswell, 1996, p. 241-244, 312-314).

Completeness The axioms above does not constitute all the required axioms for completeness. A complete axiom system can be found in section 4.6 on page 70. The axiom system presented there results in a logic denoted $\mathbf{S5}_{n,\overline{\sigma}}$. Where $\mathbf{S5}_{n,\overline{\sigma}}$ is based on the language $\mathcal{L}_{\mathsf{2QEL}}$ such that $\overline{\sigma} = \{\sigma_{OBJ}, \sigma_{LEX}\}$, $\mathbf{S5}_{n,\overline{\sigma}}$ is sound and complete with respect to the class of 2QEL models, $\mathcal{C}_{\mathsf{2QEL}}$:

Theorem 5.1 (QS5 Completeness). *The logic $S5_{n,\overline{\sigma}}$ based on language $\mathcal{L}_{n,\overline{\sigma}}$ with $\overline{\sigma} = \{\sigma_{OBJ}, \sigma_{LEX}\}$ is sound and complete with respect to the class of 2QEL models, $\mathcal{C}_{2QEL}$.*

Proof. By Theorem 4.4 and Theorem 4.5, $\mathbf{S5}_{n,\overline{\sigma}}$ is sound and complete with respect to the class $\mathsf{EQ}_{n,\overline{\sigma}}$. Soundness with respect to $\mathcal{C}_{\mathsf{2QEL}}$ follows as each $\mathcal{C}_{\mathsf{2QEL}}$ model is based on an $\mathsf{EQ}_{n,\overline{\sigma}}$ frame – i.e. everything valid in the class of frames is also valid in the class of models based on those frame. Completeness also follows from the fact that each $\mathcal{C}_{\mathsf{2QEL}}$ model is based on an $\mathsf{EQ}_{n,\overline{\sigma}}$ frame. This entails that $\Gamma \models_{\mathsf{EQ}_{n,\overline{\sigma}}} \varphi$ implies $\Gamma \models_{\mathcal{C}_{\mathsf{2QEL}}} \varphi$ for all sets of formulas Γ and formulas φ of $\mathcal{L}_{n,\overline{\sigma}}$. Hence, for same Γ and φ, $\Gamma \models_{\mathcal{C}_{\mathsf{2QEL}}} \varphi$ implies $\Gamma \vdash_{\mathbf{S5}_{n,\overline{\sigma}}} \varphi$. □

A consequence of Theorem 5.1 is that the model-theoretically specified logic $\mathsf{L}_{\mathsf{2QEL}}$ specified above and $\mathsf{QS5}_{n,\overline{\sigma}}$ are identical.

Corollary 5.1. *Where $\overline{\sigma} = \{\sigma_{OBJ}, \sigma_{LEX}\}$, $\mathsf{QS5}_{n,\overline{\sigma}} = L_{2QEL}$.*

Proof. Theorem 5.1 implies that, for any $\mathcal{L}_{\mathsf{2QEL}}$ formula φ, φ is $\mathsf{QS5}_{n,\overline{\sigma}}$ provable iff it is $\mathcal{C}_{\mathsf{2QEL}}$ valid. As $\mathsf{L}_{\mathsf{2QEL}}$ is defined as the set of $\mathcal{C}_{\mathsf{2QEL}}$ valid formulas, the result follows. □

The content of the completeness proof is that any formula can be shown to be valid in $\mathcal{C}_{\mathsf{2QEL}}$ if, and only if, it is provable in the formal logical theory. This means that the model-theoretic approach can be utilized to decide whether or not a given formula is a theorem of the theory. Hence, if the model-theoretic construction in this section can be argued to be a fair modeling of the SLC, the completeness theorem above ensures that a completely specified formal logical theory has also been found.

There has not, however, been presented any arguments to the effect that the 2QEL framework *is* a proper modeling of the SLC. That is, the model has so far not been *validated*. In the following two sections, the topic turns to validation. First, the 2QEL framework is compared to the ontologies of the SLC, and it is shown that the framework implicitly includes proper correlates of these ontologies. Second, in section 5.4, the competence types of the SLC is identified in the formal model and it is shown that the formal competence types possess the properties required to be consistent with empirical studies.

5.3 Ontology Comparison

As the reader will recall from section 2.2.3, Marconi's structure of lexical competence includes three ontologies for an agent: real-world objects, the semantic lexicon and

the two word lexicons. The first of these, the real-world objects, are easily modeled. These are represented on the semantic level by the objects of the object part of the domain, Obj, and on the syntactic level by the σ_{OBJ} terms.

Regarding the semantic lexicon and the word lexicon, then the 2QEL framework defined does include the information for modeling these, but the appropriate structures are only implicitly included. In this section, the relations and classes are defined from the indistinguishability relation in order to extract notions befitting of the semantic lexicon and the word lexicon from 2QEL models. This is done primarily model-theoretically, but the obtained classes are related to the syntactic approach as well. Their properties are discussed throughout.

The section can be seen as a validating stepping stone between the SLC and the 2QEL framework: the ontologies from Marconi's conceptual theory is correlated one-to-one with mathematical structures, and it is shown how these relate to the 2QEL framework. After this section, little reference will be made to these structures, as the relevant properties are expressible in the logic.

5.3.1 Semantic Lexicon

In the SLC the semantic lexicon consists of the agent's concepts, his mental representations of the objects that surround him, and the relations between them. As the focus here is limited to word lexicon consisting of proper names, the focus on the corresponding semantic lexicon will be on single objects. Therefore, an agent's semantic lexicon will be a set consisting of *individual concepts*. These individual concepts are defined using the main epistemological notion from the QEL framework, namely *indistinguishability*. In order to apply the notion of indistinguishability to objects rather than worlds, an *object indistinguishability relation* is defined. This is then utilized to define *individual concept classes*. These concept classes each represent an entry in the semantic lexicon, and consist of all the object indistinguishable for some agent by some feature. Finally, the semantic lexicon is defined as the set containing all such classes.

The *object indistinguishability relation* $\sim_i^{a,w}$ has three parameters: i, a and w. First, the agent i. Second, the feature by which the related objects are indistinguishable, here a, and finally the world in which this is the case, w. Where $d_1 \sim_i^{a,w} d_2$ this is read 'in state w, i cannot distinguish d_1 from d_2 by virtue of being a', for example by virtue of being 'my brother'. The relation is defined thusly:

Definition 5.9 (Object Indistinguishability Relation). Let $a \in CON$, $w \in W$ and $i \in I$. Then agent i's *object indistinguishability relation for a at w* is a binary relation $\sim_i^{a,w}$ on Obj given by

$$d \sim_i^{a,w} d' \text{ iff } \mathcal{I}(a,w) = d \text{ and } \exists w' \sim_i w : \mathcal{I}(a,w') = d'.$$

◄

An example is given after the next definition.

Though defined using $\sim_i$, the relation is not an equivalence relation, as reflexivity fails for d and $\sim_i^{a,w}$ if $\mathcal{I}(a,w) \neq d$. Still, the relation can be used to define a set of classes on the domain. These classes represent objects from the domain indistinguishable to one another by the given feature. Hence, each class consists of the objects falling under some mental heading for the given agent and therefore represents the objects in the agent's individual concept of the feature. These classes are the formal counterparts in the semantics of 2QEL to the entries in the semantic lexicon of the SLC. These classes are defined as follows:

Definition 5.10 (**Individual Concept Class**)**.** Where $\sim_i^{a,w}$ is as specified in Definition 5.9, agent i's *individual concept class for a at w* is defined by

$$C_i^{a,w}(d) = \{d' : d \sim_i^{a,w} d'\}.$$

◀

The set $C_i^{a,w}(d)$ consists of the objects indistinguishable to agent i via feature a from object d in the part of the given model connected to w by $\sim_i$.

To exemplify these two definitions in use, consider again the case with the two boxes and the cat, illustrated in Figure 3.2.2 on page 43. Again let constants g, w and c denote the grey box, the white box and the box containing the cat, respectively. In the illustrated model, there are two worlds, w_1 and w_2, related by $\sim_i$. Assume w_1 is the actual world. As agent i can identify the grey box, the interpretation of g is constant across these worlds, i.e. $\mathcal{I}(g,w_1) = \mathcal{I}(g,w_2) = d_1$. From this it follows that the only object related to d_1 by $\sim_i^{g,w_1}$ is d_1 itself. To see this, notice first that $d_1 \sim_i^{g,w_1} d_1$. This follows as $\mathcal{I}(g,w_1) = d_1$, since this satisfies the existential claim of Definition 5.9. Second, there can be no other objects related by $\sim_i^{g,w_1}$ to d_1 as no further worlds exist to satisfy the existential claim. Hence, the only object that i cannot distinguish from the actual grey box, d_1, by virtue of being the grey box, g, is the grey box itself, d_1. Therefore, agent i's individual concept class for the grey box, g, contains only the grey box itself, d_1. That is, $C_i^{g,w_1}(d_1) = \{d_1\}$.

As opposed to the grey box, the agent is uncertain regarding which of the two objects is the box containing the cat. In w_1, the box containing the cat, c, is d_2, i.e. $\mathcal{I}(c,w_1) = d_2$. In w_2, it is d_1, i.e. $\mathcal{I}(c,w_2) = d_1$. Utilizing Definition 5.9, it can be seen that $d_2 \sim_i^{c,w_1} d_2$ and $d_2 \sim_i^{c,w_1} d_1$. The first case follows by the same argument as above. The second follows as w_2 satisfies the existential claim for d_1 in that $\mathcal{I}(c,w_2) = d_1$. Hence, the objects that i cannot tell apart from the box actually containing the cat, d_2, by virtue of being the box containing the cat, c, is the white box, d_2, and the grey box, d_1. Put differently, $C_i^{c,w_1}(d_2) = \{d_1, d_2\}$.

In Figure 3.2.2 on page 43 the agent's individual concept classes for the grey box, the white box and the box concealing the cat are illustrated as two singletons, and one set containing both the grey and the white box.

Based on the definition of individual concept classes, the semantic lexicon for agent i may now be defined as follows:

Definition 5.11 (**Semantic Lexicon**)**.** The semantic lexicon of agent i in the pointed 2QEL model (M, w) with $d \in Obj$ and $a \in CON$ is the set

$$\mathsf{SL}_i = \{C_i^{a,w}(d) : C_i^{a,w}(d) \neq \emptyset\}.$$

◀

The requirement that the empty set is not included in the semantic lexicon reflects the idea that there are no empty thoughts. It is included in order to eliminate nonsense concepts.

In the example above, it would be the case that

$$\mathsf{SL}_i = \{C_i^{g,w_1}(d_1), C_i^{w,w_1}(d_2), C_i^{c,w_1}(d_1), C_i^{c,w_1}(d_2)\}$$
$$= \{\{d_1\}, \{d_2\}, \{d_1, d_2\}\}$$

The object indistinguishability relations and the individual concept classes properly reflect the notion of identifiability. The key aspect of an agent's ability to identify an object is that the object is *singled out*, in the sense that the object is distinguishable from all other objects by the agent. In order for the agent to identify an object, the agent must have an *unambiguous* concept of the object. Proposition 5.1 nicely ties together the two ideas of an unambiguous individual concept and the notion of identifiability, by stating that the given object is distinguishable from all other objects by the object indistinguishability relation if, and only if, the agent is able to identify the given object.

Proposition 5.1. *An object can be identified if, and only if, the appropriate concept class has cardinality 1, i.e.*

$$M, w \models_v \exists x K_i(x = a) \tag{5.3.1}$$

if, and only if,
$$|C_i^{a,w}(d)| = 1$$

Proof. (5.3.1) holds iff $\mathcal{I}$ is constant for a across all w' such that $w \sim_i w'$. This is the case iff d is the only object such that $d \sim_i^{a,w} d$, which again is the case iff the class $C_i^{a,w}$ is a singleton. □

The formally defined semantic lexicon nicely represents the semantic lexicon of the SLC. Each entry in the semantic lexicon reflects the agent's concept of an object and the agent's information regarding the identity of the object. Proposition 5.1 shows that this information exactly corresponds to the criteria for identifiability from the QEL framework. The proposition relates the semantic lexicon and an agent's ability regarding the real-world objects in a clear way. It therefore provides a clear criteria for the second part of the two-stage process of application to be successful: the agent must possess a singleton concept class of the meaning of the name in question in order to identify the referent. This will be returned to in section 5.4 below.

Syntactical representation   In order for the syntactically represented logic to also include the semantic lexicon, the concept classes must somehow be expressible in the logic. A suitable notion can be defined by using a *modal predicate*. In order to define the predicate, the following observation is required:

Proposition 5.2.

$$M, w \models_v P_i(a = b) \tag{5.3.2}$$

if, and only if,

$$\mathcal{I}(b, w') \in C_i^{a,w}(\mathcal{I}(a, w))$$

for all w' such that $w \sim_i w'$ and $M, w' \models_v (a = b)$.

Proof. Initially, $M, w \models_v P_i(a = b)$ holds only if $M, w' \models_v (a = b)$ for some w' such that $w \sim_i w'$. For such w', the identity $\mathcal{I}(a, w') = \mathcal{I}(b, w')$ holds. By the definition of $\sim_i^{a,w}$, it follows that $\mathcal{I}(a, w) \sim_i^{a,w} \mathcal{I}(a, w')$, so $\mathcal{I}(a, w) \sim_i^{a,w} \mathcal{I}(b, w')$ by the shown identity. By the definition of $C_i^{a,w}(\mathcal{I}(a, w))$ it follows that $\mathcal{I}(b, w') \in C_i^{a,w}(\mathcal{I}(a, w))$.

As all relationships used are bi-directional, the proposition is shown. □

It follows from Proposition 5.2 that it is possible to define a modal predicate for each constant a and each agent i with an extension equal to that of the agent's concept of a at w:

Definition 5.12 (Individual Concept Predicate). Let (M, w) be a pointed 2QEL model, v a valuation. Then, where $a, b \in CON$ and $i \in I$, define agent i's *individual concept predicate for a* by

$$M, w \models_v C_i^a(b) \leftrightarrow P_i(a = b)$$

◀

The predicate C_i^a captures agent i's individual concept of a at w. The predicate is called "modal" as it is defined using a modal notion and the use of classic existential generalization does not preserve truth when the predicate occurs in the antecedent formula. For example does $M, w \models_v K_i \neg(a = b)$ (equivalent to $M, w \models_v \neg P_i(a = b)$) not imply $M, w \models_v \exists x K_i \neg(x = b)$.

The definability of such individual concept predicates shows that the semantic lexicon aspect of the SLC is appropriately represented in the logical theory.

Properties of the semantic lexicon   The way things have been defined ensures that the system does not exhibit an unwanted collapse between concepts. Consider again the example of the cat and the boxes. In this case, agent i considers it possible that both the grey and the white box is the box that contains the cat, i.e.

$$M, w_1 \models_v P_i(g = c) \land P_i(w = c)$$

Therefore the agent's concepts of g and w have a non-empty intersection as

$$\mathcal{I}(c, w_1) \in C_i^{g,w_1}(d_1) \text{ and } \mathcal{I}(c, w_1) \in C_i^{w,w_1}(d_2)$$

by Proposition 5.2. Though the concepts overlap, this does not imply that the classes are identical. Such a collapse of concepts would be unwanted, as the agent is able to distinguish between the grey box and the white box, which in turn means they should have distinct entries in the semantic lexicon. That overlapping concepts do not imply the identity between such is the content of the following proposition.

Proposition 5.3.

$$M, w \models_v P_i(a = b)$$

does not imply that

$$C_i^{a,w}(d) = C_i^{b,w}(d)$$

Proof. First, note that

$$P_i(a = b) \land P_i(b = c) \land \neg P_i(a = c) \tag{5.3.3}$$

is a contingent formula, and hence satisfiable. Assume model M satisfies (5.3.3) at w and that w is the required state satisfying $(a = b)$, so for some $d \in Obj$, $\mathcal{I}(a, w) = \mathcal{I}(b, w) = d$.

From the first conjunct in (5.3.3) and Proposition 5.2 it follows that $\mathcal{I}(b, w) \in C_i^{a,w}(d)$ and $\mathcal{I}(a, w) \in C_i^{b,w}(d)$. The second conjunct and Proposition 5.2 implies that $\mathcal{I}(c, w') \in C_i^{b,w}(d)$, for some w' such that $w \sim_i w'$. From the assumption that $C_i^{a,w}(d) = C_i^{b,w}(d)$, a contradiction may now be derived. Since this assumption implies that $\mathcal{I}(c, w') \in C_i^{a,w}(d)$. But by Proposition 5.2, this entails that $M, w \models_v P_i(a = c)$, contrary to the assumption of (5.3.3). $\square$

Where Proposition 5.2 relates the P_i operator to object inclusion in concept classes, the dual operator K_i can be related to the identity of concepts. In particular, if an agent knows that two objects are identical, the agent's individual concepts of the objects will be the same.

Proposition 5.4. *Knowledge of identity implies identity of concepts.*

$$M, w \models_v K_i(a = b)$$

implies

$$C_i^{a,w}(\mathcal{I}(a, w)) = C_i^{b,w}(\mathcal{I}(b, w))$$

Proof. Assume that $M, w \models_v K_i(a = b)$. Then $M, w \models_v a = b$, and hence $\mathcal{I}(a, w) = \mathcal{I}(b, w)$, so $\mathcal{I}(b, w) \in C_i^{a,w}(\mathcal{I}(a, w))$. Now regard some arbitrary $d \in C_i^{a,w}(\mathcal{I}(a, w))$. It is shown that $d \in C_i^{b,w}(\mathcal{I}(b, w))$. This is the case iff $\mathcal{I}(b, w) \sim_i^{b,w} d$, by the definition of concept classes. This is again the case iff $\exists w' \sim_i w : \mathcal{I}(b, w) = \mathcal{I}(b, w)$ and $\mathcal{I}(b, w') = d$, by the definition of the object indistinguishability relation. From the assumption that $d \in C_i^{a,w}(\mathcal{I}(a, w))$, it follows that $\exists w' \sim_i w : \mathcal{I}(a, w') = d$. As $\mathcal{I}(a, w'') = \mathcal{I}(b, w'')$ for all $w'' \sim_i w$ by the assumption of $M, w \models_v K_i(a = b)$, it follows that $\exists w' \sim_i w : \mathcal{I}(b.w') = d$. Hence $\mathcal{I}(b, w) \sim_i^{b,w} = d$. So $d \in C_i^{b,w}(\mathcal{I}(b, w))$, why $C_i^{a,w}(\mathcal{I}(a, w)) \subseteq C_i^{b,w}(\mathcal{I}(b, w))$. The opposite inclusion is symmetric and therefore omitted. $\square$

Notice that the implication in the above is not bi-directional. That the opposite implication does not hold is a wanted feature of the system since the agent may have completely overlapping concepts because the agent possesses too little information. Again returning to the cat in the box example, consider the case where i is blind-folded. In this case, i cannot distinguish *any* of the boxes from each other[8], and all three individual concepts therefore completely overlap. In this case, the agent should not know that the grey and the white box are the same. This example provides the intuition behind the proof for the following proposition:

Proposition 5.5. *Identity of concepts does not imply knowledge of identity.*

$$C_i^{a,w}(\mathcal{I}(a,w)) = C_i^{b,w}(\mathcal{I}(b,w))$$

does not imply

$$M, w \models_v K_i(a = b)$$

Proof. Assume that $C_i^{a,w}(\mathcal{I}(a,w)) = C_i^{b,w}(\mathcal{I}(b,w))$ and that $d \in C_i^{a,w}(\mathcal{I}(a,w)) = C_i^{b,w}(\mathcal{I}(b,w))$. Then $\exists w' \sim_i w : \mathcal{I}(a,w') = d$ and $\exists w'' \sim_i w : \mathcal{I}(b,w'') = d$. For $M, w \models_v K_i(a = b)$, it must be the case that $\mathcal{I}(a,w^*) = \mathcal{I}(b,w^*)$ for all $w^* \sim_i w$. A counterexample to this is provided by the following model. Let $M = \langle W, (\sim_i)_{i \in I}, \mathcal{I} \rangle$, where $W = \{w, w'\}$ and $w \sim_i w'$, with the interpretation of constants a and b as given by the following table:

	w	w'
a	d	d'
b	d'	d

Here, $C_i^{a,w}(\mathcal{I}(a,w)) = \{d, d'\} = C_i^{b,w}(\mathcal{I}(b,w))$, but $M, w \models_v \neg K_i(a = b)$. ☐

5.3.2 Word Lexicon

The word lexicon is represented in a similar way to the semantic lexicon. On the semantic side, the word lexicon of agent i is represented by classes on the name part of the domain, Nam. Each such class will be a singleton set, by the assumption of syntactic competence. This precisely reflects the assumption that each agent is able to distinguish between all name types, i.e. have an unambiguous individual concept of each type. These equivalence classes may be obtained by the following definition:

Definition 5.13 (Name Class). Where $\dot{n} \in Nam$ and $n \in LEX$, agent i's *name class for n* is

$$C_i^n(\dot{n}) = \{\dot{n}' : \mathcal{I}(n) = \dot{n}'\}$$

◄

---

[8]By virtue of being grey, white or containing the cat.

The definition is simpler as the interpretation of names is not world relative. That the defined classes exhibit the intended properties is captured by the following proposition.

Proposition 5.6. *For all $\dot{n} \in Nam$, $C_i^n(\dot{n})$ form a singleton equivalence class and the set $\{C_i^n(\dot{n}) : \dot{n} \in Nam\}$ form a partition on Nam.*

Proof. First, to see that $\cup_{\dot{n} \in Nam} C_i^n(\dot{n}) = Nam$, recall that for all $\dot{n} \in Nam$ there is some $n \in LEX$ such that $\mathcal{I}(n) = \dot{n}$ by the definition of $\mathcal{I}$. So each name is an element of a class. As all $\dot{n} \in Nam$ are in some class, the union of these classes contain all $\dot{n} \in Nam$. Secondly, that $C_i^n(\dot{n})$ is a singleton for each $\dot{n} \in Nam$ can be argued as follows. Assume that $\dot{n}, \dot{n}' \in C_i^n(\dot{n})$. Then, for some $n' \in LEX$, $\mathcal{I}(n') = \dot{n}'$. So $\mathcal{I}(n') \in C_i^n(\dot{n})$. Hence, by the definition of $C_i^n(\dot{n})$, it follows that $\mathcal{I}(n') = \dot{n}$, and it can be concluded that $\dot{n}' = \dot{n}$, why $C_i^n(\dot{n})$ is a singleton set. Singletons are obviously either disjunct or identical, why the statement is proven. $\square$

That all name concepts for any agent are singleton sets reflects the validity of the KNI axiom for lexical terms, as mentioned in section 5.2.2. The KNI axiom can be re-written as

$$(t \neq t') \rightarrow \neg P_i(t = t').$$

The axiom states that if two name tokens are not instances of the same name type, these names will *not* be in the same concept class. The framework thus nicely captures that the agents are syntactically competent with respect to all words in Nam. Again, this is an aspect not discussed by Marconi.

A modal predicate with an extension equal to the concept class of a name type can easily be defined using the same approach as for the semantic lexicon. As it will not be used, the definition is omitted.

It should be noted that the assumption is false for certain groups of human agents – an example could be people with impaired hearing, who may not be able to identify which name was uttered in a given context. In order to model such lacking *phonetic* competence, one should allow for possible falsity of KNI and hence opt for a contingent identity system of names.

From Definition 5.13, the word lexicon for agent i is defined as the set of all word classes:

Definition 5.14 (**Word Lexicon**)**.** The *word lexicon* of agent i in the pointed 2QEL model (M, w) with $\dot{n} \in Nam$ is the set

$$\mathsf{WL}_i = \{C_i^n(\dot{n}) : n \in LEX\}.$$

◀

5.3.3 Closing Remarks

The formal version of the semantic lexicon, SL_i, and the word lexicon, WL_i, along with their constituting classes will not be used much in any explicit form. The individual concept classes and the name class constituting SL_i and WL_i were defined

from the indistinguishability relation from the 2QEL models. Further, the properties most interesting for the present work, such as concept inclusion and object identification, were seen to be expressible in the language $\mathcal{L}_{2QEL}$. Hence, explicit reference to these classes and the lexica are superfluous: what is relevant from these notions can readily be expressed in the 2QEL framework. The formal lexica were introduced to validate ontologies of the 2QEL framework as a formal modeling of the SLC ontologies. This has been done. It has been shown in this section that appropriate notions of individual concept classes and name classes constituting the semantic and word lexica can be extracted from the 2QEL framework. The section can in this light be seen as a stepping stone between the SLC and 2QEL, showing how the SLC notions can be expressed using the 2QEL framework. As the relevant features are readily expressible without making explicit reference, this stepping stone will not be referred much to in the ensuing chapters.

5.4 Competence Types

Having identified the three ontologies present in Marconi's theory, the competence types over these may now be given a formal reading. In addition to referential and inferential competence, an extra 'competence type' is present in the epistemic logical framework, namely 'worldly competence'. The aspects of this 'competence' are exactly those usually modeled in quantified epistemic logic and has as such been expounded in chapter 3.

5.4.1 Referential Competence

Regarding referential competence, recall that this compromises two distinct subsystems between names and objects, relating these through the semantic lexicon. The two relations are *application* and *naming*. Regarding application, then an agent can apply a name if when presented with a token of the name, the agent can identify the appropriate referent. This ability can be defined as follows:

Definition 5.15 (**Application**). Agent i can *apply* name n in pointed model (M, w) iff

$$M, w \models_v \exists x K_i(\mu(n) = x) \tag{5.4.1}$$

◀

Recalling the reading of *de re* formulas (cf. (3.2.1) on page 44), the definition literally states that agent i can identify the meaning/referent of n.

The second referential competence type is naming. To be able to name an object, agent i is required to be able to produce a name when presented with an object, say a. This can equally well be expressed using the same *de re*-type formula as in the definition above:

Definition 5.16 (**Naming**). Agent i can *name a* in pointed model (M, w) iff

$$M, w \models_v \exists \dot{x} K_i(\mu(\dot{x}) = a) \tag{5.4.2}$$

◀

Here, the definition states that the agent can identify a name the referent of which is a.

As mentioned in section 2.2.3, application and naming are distinct subsystems in the sense that an agent may loose the ability to name an object while maintaining the ability to apply a name referring to the object or *vice versa*. This feature is preserved in the 2QEL framework as neither (5.4.1) nor (5.4.2) imply one another. This is shown in section 5.4.3 below.

5.4.2 Inferential Competence

Regarding inferential competence, the present framework is rather limited in the features expressible. This is a direct consequence of the simplified version of the SLC modeled. In particular, the choice of only including proper names in the word lexicon limits the types of inferential competence to knowing relations between referring names and not inferential knowledge regarding names and verbs. As an example, knowledge of true subject-predicate sentences such as 'a is P' cannot be expressed as the word lexicon does not contain 'is' nor 'P', but only a name for a.

The key notion of inferential competence modeled is therefore knowledge of the co-reference of proper names:

Definition 5.17 (**Knowledge of Co-reference**). Agent i *knows that names n and n' co-refer* in pointed model (M, w) iff

$$M, w \models_v K_i(\mu(n) = \mu(n')) \tag{5.4.3}$$

◀

This type of inferential competence satisfies the requirement of relating entries in the word lexicon through the semantic lexicon: where $\mathcal{I}(n)$ and $\mathcal{I}(n')$ are elements in two name classes from the word lexicon, $\mathcal{I}(\mu(n), w)$ and $\mathcal{I}(\mu(n'), w)$ are elements from individual concept classes in the semantic lexicon. Knowledge of co-reference implies by (Proposition 5.4) that these concept classes are identical, which then facilitates the connection between the meaning of the two names for the agent.

Knowledge of co-reference is a special case of full inferential competence. When the agent has knowledge of the co-reference of two names, this does not mean that he is inferentially competent in the terminology of Marconi. Marconi's use is holistic, as the agent has to be able to supply a variety of synonyms and properties to be deemed inferentially competent with respect to a name. Such a holistic notion of inferential competence is defined as follows:

Definition 5.18 (**Full Inferential Competence**)**.** Agent i is *fully inferentially competent with respect to n* in pointed model (M, w) iff

$$M, w \models_v (\mu(n) = \mu(n')) \rightarrow K_i(\mu(n) = \mu(n')) \tag{5.4.4}$$

for all $n' \in LEX$. ◀

that is, the agent is fully inferentially competent with n iff the agent knows of *every name co-referring with n* that these two co-refer. The agent may, of course, be less competent and know only of the co-reference of a strict subset of a set of synonyms. This will be equivalent to the agent having knowledge of co-reference of the names in the subset.

5.4.3 Dissociation of Competence Types

The three types of competence identified in (Marconi, 1997) were there found to be dissociate. For the modeling to be able to capture the possible cases reported, it will thus be required that the corresponding formal types of competence identified preserve this dissociation. Accordingly, neither of (5.4.1), (5.4.2) or (5.4.3) should imply one another. This is the content of the following four propositions.

Proposition 5.7. *Application does not imply naming*

Proof. Regard a model in which $W = \{w, w'\}$, $w \sim_i w'$ and

$$\mathcal{I}(\mu(n), w) = \mathcal{I}(\mu(n), w') = d$$

Then $M, w \models_v \exists x K_i(\mu(n) = x)$. That is, i can apply n.

Assume further that $\mathcal{I}(a, w) = d$ and $\mathcal{I}(a, w') = d'$. Then, though $M, w \models_v (\mu(n) = a)$, it is still the case that $M, w \models_v \neg\exists \dot{x} K_i(\mu(\dot{x}) = a))$. Hence the agent cannot name the referent of n. □

Proposition 5.8. *Application does not imply knowledge of co-reference*

Proof. To see that (5.4.1) does not imply (5.4.3), regard the model from proposition Proposition 5.7. In this model, the agent can apply n. Make the additional assumption that

$$\mathcal{I}(\mu(n), w) = \mathcal{I}(\mu(n'), w)$$

to the effect that $M, w \models_v (\mu(n) = \mu(n'))$. Finally, assume that $\mathcal{I}(\mu(n'), w') = d''$.

Then $M, w \models_v \neg K_i(\mu(n) = \mu(n'))$. The agent is still able to apply n, but does not know that n and n' in fact co-refer. □

Proposition 5.9. *Naming does not imply application nor inferential competence*

Proof. A new counter-example is constructed to show that (5.4.2) does not imply (5.4.1) nor (5.4.3). Again, let $W = \{w, w'\}$ with $w \sim_i w'$. Let

$$\mathcal{I}\left(\mu(\dot{x}), w\right) = \mathcal{I}\left(a, w\right) = d \text{ and } \mathcal{I}\left(\mu(\dot{x}), w'\right) = \mathcal{I}\left(a, w'\right) = d'$$

Here, (5.4.2) holds at w, whereas (5.4.1) does not.

To make (5.4.3) false at w, assume that $v(\dot{x}) = \mathcal{I}(n)$ and for a further name n' that

$$\mathcal{I}\left(\mu(n'), w\right) = \mathcal{I}\left(\mu(n), w\right) \text{ and } \mathcal{I}\left(\mu(n'), w'\right) = d$$

This provides a counter-example. $\square$

Proposition 5.10. *Inferential competence implies neither application nor naming*

Proof. To see that (5.4.3) implies neither (5.4.1) nor (5.4.2), let

$$\begin{aligned}
d =& \mathcal{I}\left(a, w\right) = \mathcal{I}\left(\mu(n), w\right) = \mathcal{I}\left(\mu(n'), w\right) \text{ and} \\
d' =& \mathcal{I}\left(\mu(n), w'\right) = \mathcal{I}\left(\mu(n'), w'\right) \text{ and} \\
d'' =& \mathcal{I}\left(a, w'\right).
\end{aligned}$$

Then (5.4.3) is satisfied at w, whereas (5.4.1) and (5.4.2) are both false. $\square$

5.5 Further Properties

An advantage of formal theories over informal ones is that it is quite easy to deduce theorems and properties of the formal theory. Investigating such theorems and properties can in turn be used to discover otherwise overlooked features of the conceptual theory, provide an overview of the theory and its implicit assumptions and yield hypotheses which can be evaluated using new empirical data. This section is dedicated to these three ventures.

In relation to semantic competence, the most interesting features obviously regard the knowledge agents have of words and their meanings. In order to provide an albeit limited picture of the features of the formal theory, a selection of formulas capturing interesting properties are now presented. As a first example, the theory entails that any agent knows that every name has a referent. This is captured by

$$K_i \forall \dot{x} \exists x (\mu(\dot{x}) = x) \tag{5.5.1}$$

The validity of (5.5.1) stems from the assumption that μ is a total function. To work with systems with non-denoting names, μ should be defined as a partial function, and the axiom system should be modified. As (5.5.1) is a $\mathcal{C}_{2\mathsf{QEL}}$ validity, it is $\mathsf{L}_{2\mathsf{QEL}}$-provable.

Furthermore, the agents also know that the names of *LEX* are unambiguous. This is reflected by the provability of

$$\forall x \forall y \forall \dot{x} K_i ((\mu(\dot{x}) = x) \wedge (\mu(\dot{x}) = y) \rightarrow (x = y))$$

which captures that the agents know that the names of *LEX* are assigned unambiguous meaning by μ. Put differently, the agents know that the reference relation is a function.

That μ was not assumed surjective results in the invalidity of

$$K_i \forall x \exists \dot{x} (\mu(\dot{x}) = x) \tag{5.5.2}$$

stating that agent i knows of every object that it is named. Only in the models of $\mathcal{C}_{2\mathsf{QEL}}$ where μ is surjective will (5.5.2) hold. Nothing prevents such models from being constructed, and hence (5.5.2) is $\mathcal{C}_{2\mathsf{QEL}}$-satisfiable. It is therefore also $\mathsf{L}_{2\mathsf{QEL}}$-consistent.

Though the validity of (5.5.1) ensures that i knows that all names refer, it is not assumed that the agent knows *what they refer to*. Unless specifically assumed, this knowledge is not the case, which is represented by

$$\forall \dot{x} \exists x K_i (\mu(\dot{x}) = x) \tag{5.5.3}$$

being invalid in $\mathcal{C}_{2\mathsf{QEL}}$. This is natural as most competence types are made as substantial assumptions: the theory should not entail that the agents are always able to apply all names, as this would be inconsistent with empirical findings. Depending on the agents, word lexicon and the object domain under consideration, it can sometimes be a natural assumption: when playing cards, for example, all players can both apply the names of and name all cards in the deck. This possibility is consistent with the theory, in that the formula (5.5.3) is $\mathcal{C}_{2\mathsf{QEL}}$-satisfiable, why it is also $\mathsf{L}_{2\mathsf{QEL}}$-consistent.

In the class of $\mathcal{C}_{2\mathsf{QEL}}$ models, it is possible to satisfy

$$P_i((n \neq n') \wedge \mu(n) = \mu(n'))$$

which states that even though two names are instances of different name types, agents may still consider it possible that they co-refer. Hence, the system allows for the *arbitrariness of sign*, i.e. that there is nothing essential binding a given sign to the object it names.

5.5.1 Further Competence Types

In (Marconi, 1997), there is not much discussion about different strengths of competence. In many ways, the conceptual model is less precise than the formal theory presented here. The difference between knowledge of co-reference (Definition 5.17) and full inferential competence (Definition 5.18), for example, is clear in the formal framework, but is not explicitly discussed by Marconi. Further, the type of meta-competence encoded in(5.5.1) or the universal application of (5.5.3) is not mentioned.

One very interesting property of the formal theory is an additional competence type, which is $\mathsf{L}_{2\mathsf{QEL}}$-provable of any agent. This is *trivial competence*:

Definition 5.19 (**Trivial Competence**). An agent i is said to be *trivially competent with n* iff

$$K_i(\mu(n) = \mu(n)) \tag{5.5.4}$$

◀

Any agent is trivially competent with any name, as (5.5.4) is a $\mathcal{C}_{\mathsf{2QEL}}$ validity.

Thought trivial competence may seem uninteresting and trivial, the provability of (5.5.4) captures an important implicit assumption of the $\mathsf{2QEL}$ framework. What is implied by the provability of trivial competence is that any agent *has an individual concept for the meaning of every name*. Stated otherwise, for every name $n \in LEX$, the class $C_a^{\mu(n),w}(\mathcal{I}(\mu(n), w))$ is a member of i's semantic lexicon.

As the agents are assumed to be syntactically competent and know that all names refer, this is not an unnatural property. Still, it is not trivial, for it ensures that there is a connection between the word lexicon and the semantic lexicon *for every name*. It may be that the entry in the semantic lexicon encodes *no information*, since the concept may contain the entire object domain, but the connection will still exist.

The provability of trivial competence may seem natural if presented as follows: (5.5.4) states that an agent always identifies the meaning of a name with the meaning of the same name, possibly without knowing anything about this meaning. This may be seen as what allows children to correctly, but trivially, answer many questions. In reply to 'who was Napoleon?', the trivial answer 'Napoleon' is not wrong.

A further competence type identifiable in the formal framework is what may be called *correlation*. This type results when the agent is able to correlate a name with an non-linguistic entry in the semantic lexicon, but where the latter is not assumed to be an unambiguous concept:

Definition 5.20 (**Correlation**). Agent i *correlates name n with a* in pointed model (M, w) iff

$$M, w \models_v K_i(\mu(n) = a) \tag{5.5.5}$$

◀

The definition states that i knows that the referent of n is the object a, yet it does not imply that i is able to identify this object. In the running example, introduce the name *Schrödinger's Box* for the box containing the cat. Then i would know that $\mu(\textit{Schrödinger's Box})$ was object c, without being able to identify either.

This ability is referred to as *correlation* since the agent correlates meaning with concept, but possibly not concept with world. By Proposition 5.4, (5.5.5) implies that agent i's individual concepts of $\mu(n)$ and a are identical. Correlation is thus a relation between word lexicon and semantic lexicon, that does not imply that the agent can apply the name. Whether correlation should be classified as a referential or an inferential competence type does not seem important.

Both trivial competence and correlation will be used in the analyses of chapter 6.

5.5.2 Implications between Competence Types

As shown above, application, naming and inferential competence do not entail one another. The concepts are closely related, and in various conjunctions, implicational relationships arise. Of particular interest later will be those between application and inferential competence. The following four facts capture such relationships that will be used later on in the analyses of Frege's puzzles.[9]

The first fact shows that the formal theory entails that agents are able to reason about application and co-reference. To exemplify, imagine an agent able to apply two co-referring names in parallel: the agent can point to the referent of n with his left hand and to the (same) referent of n' with his right – Fact 5.1 states that the agent should then know that the two names co-refer:

Fact 5.1. *Application of co-referring names implies knowledge of co-reference, i.e.*

$$If\ M, w \models_v (\mu(n) = \mu(n'))\ then$$
$$M, w \models_v \exists x K_i(\mu(n) = x) \land \exists y K_i(\mu(n') = y) \to K_i(\mu(n) = \mu(n'))$$

This seems to be a small, natural deductive step for human agents, but whether it is consistent with empirical evidence is not discussed by Marconi.

A similar small deductive step is involved in the following: in case agent i is able to apply n, then for any name n' for which he has knowledge of co-reference with n, the theory states that he should be able to apply n':

Fact 5.2. *Application and inferential competence implies application. That is,*

$$M, w \models_v (\mu(n) = \mu(n'))\ implies$$
$$M, w \models_v \exists x K_i(\mu(n) = x) \land K_i(\mu(n) = \mu(n')) \to \exists y K_i(\mu(n') = y)$$

As mentioned in section section 5.4, a weak competence type not discussed in (Marconi, 1997), namely correlation, can be identified in the formal framework. The competence type can be seen as being weaker than application as the latter implies the former, by the following fact.

Fact 5.3. *Application implies correlation, i.e.*

$$\mathcal{C}_{2QEL} \models \exists x K_i(\mu(n) = x) \to K_i(\mu(n) = a)$$

The converse implication does not hold, as the agent will lack identificatory knowledge. But in case agent i has an unambiguous concept of a, then this in conjunction with the correlation of this concept with the meaning of a name does imply the ability to apply the name:

Fact 5.4. *Unambiguous concept and correlation implies application, i.e*

$$\exists x K_i(x = a) \land K_i(\mu(n) = a) \to \exists y K_i(\mu(n) = y)$$

---

[9]Proofs of the facts are not included. In all cases, the assumption of the negated fact will soon be seen leading to a contradiction.

That is, the theory states that agents able to identify an object while also having a name associated with the individual concept of the object should be able to apply the name to the object.

5.5.3 Validation and Deductive Skills

The four facts above are examples of testable predictions of the formal theory. Finding subjects with the correct abilities and testing the semantic skills can be used to evaluate the theory. However, problems arise if contradicting evidence is found.

One problem regarding the present modeling is the *Problem of Logical Omniscience*, first discussed by Hintikka (1962). The axiom system and semantics of QEL ensure that any agent knows all logical consequences of his knowledge. As the axioms include the axioms of propositional logic, the agent automatically and instantaneously knows all of the infinite set of theorems of propositional logic. This is obviously not the case for any, finite human agent. As the 2QEL framework is an extension of QEL, the problem carries over. Hence, the proposed framework *cannot* be a precise account of *any* subject's knowledge.

If the problem of logical omniscience is ignored or solved, a problem regarding the subject's deductive skills in relation to possibly falsifying evidence still remains. Consider, for example, a subject satisfying the antecedent of Fact 5.2 above by being able to identify the referent of n and presents n' when asked to supply a synonym. In this case, the theory states that the subject should be able to identify the referent of n', too. But suppose the subject cannot do this. Then either the theory ought to be rejected, or it can be salvaged by excuses regarding the deductive skills of the subject. It might be argued that the subject is tired or that the subject has too short an attention span to complete the basic deduction required to obtain knowledge of the consequent. This way, the blame can be shifted from the theory of semantic competence to the underlying theory of deductive skills.

If extensions of epistemic logic are used, then it is already known from the problem of logical omniscience that these do not correctly represent human cognitive skills. But if a logic without logical omniscience is used, it may *still* be argued that the problem lies with this underlying theory of deduction, rather than the theory of semantic competence. Hence, in order to be able to properly validate the present framework, the theory of semantic competence should ideally be constructed in a system *already shown to properly represent human deductive skills*. Until such a system is constructed and amply validated, the outlook for validating the theory constructed here is bleak.

5.6 Conclusions

In this chapter, a formal logical theory for a simplified version of Marconi's structure of lexical competence has been constructed. This formal theory has been identified with the logic $\mathsf{QS5}_{n,\overline{\sigma}}$ with $\overline{\sigma} = \{\sigma_{OBJ}, \sigma_{LEX}\}$. The logic was obtained via a com-

pleteness result, showing that this logic was sound and complete with respect to the class of models considered. This class of models have in turn been argued as fairly representing the important aspects from Marconi's theory by i) correctly representing the ontological elements and ii) preserving the implicational relationships between the competence types. Finally, the proposed model was investigated with respect to some of the implicit assumptions made, some novel competence types were identified, some facts about the theory was listed and a problem regarding validation was discussed.

In the following chapter, the theory of semantic competence devised here will be applied to Frege's puzzles, and it will be shown that the theory sheds new light on these problems.

Appendix: Towards a Hierarchy of Semantic Competence Types

Note: This section is an appendix to the above, and the content *is not used* in the remainder of the thesis. The reader can feel free to skip this section.

Given the implicational relationships between conjunctions of different competence types noted in section 5.5.2, one may be led to seek out a general hierarchy of both the types and strengths of semantic competence. Constructing a partially ordered hierarchy could lead to an exact taxonomy of all competence types possible in the presented framework. Alas, due to the many possible constellations of competence types, names and objects that must possibly be considered, such a hierarchy becomes quite complex. A complete elucidation of such a hierarchy is beyond this section, and only a tentative exposure will therefore be given. In particular, a rudimentary idea for a possible method of construction will be presented along with the sizes of the corresponding hierarchies.

In order to construct such a hierarchy, regard the set of name tokens, LEX. This set is partitioned by the identity into a finite partition. The partition is finite as the identity is defined relative to the interpretation $\mathcal{I}$ and the finite set Nam. The partition is defined as follows:

Definition 5.21 (Name Type Partition).

$$[LEX] = \{[\dot{n}_1], [\dot{n}_2], ..., [\dot{n}_k]\}$$

where $[\dot{n}_i] = \{n \in LEX : \mathcal{I}(n) = \dot{n}_i\}$. ◀

Each cell $[\dot{n}_i]$ hence consists of the name tokens $n \in LEX$ which are mapped to $\dot{n}_i$ by $\mathcal{I}$. The partition $[LEX]$ thus consisting of name sets of name tokens, pooled by being of the same type. They will therefore be denoted *name types* in the following. No confusion should arise between this way of regarding name types and the name types of Nam.

The starting point of the construction of the hierarchy will be sets of co-referring name types:

Definition 5.22 (Co-reference and Synonyms). Let $M \in \mathcal{C}_{2\mathsf{QEL}}$ with actual world w. Two name tokens n and n' are said to be *co-referring under* μ if

$$\mathcal{I}\left(\mu(n), w\right) = \mathcal{I}\left(\mu(n'), w\right).$$

If $n \in [\dot{n}_i]$ and $n' \in [\dot{n}_j]$, the name types $[\dot{n}_i]$ and $[\dot{n}_j]$ are said to be *co-referring* or to be *synonyms*. ◄

Regard a set of co-referring name types N from a name type partition $[LEX]$, i.e. $N \subseteq [LEX]$. Then each possible partition of N can be regarded as representing a way of being inferentially competent with the set of synonyms since each partition may be seen as induced by the agent's knowledge of the names' co-reference. To see this, define first the set of *known synonyms of n*:

Definition 5.23 (Known Synonyms). For each $n \in LEX$, let the set of *known synonyms of n* be
$$\bar{n} = \{n' : M, w \models_v K_i(\mu(n) = \mu(n'))\}$$

The sets $\bar{n}$ are termed *KS sets.* The set of all KS sets is denoted KS. ◄

The set $\bar{n}$ is the set of all name tokens which the agent knows to be co-referring with n. KS can be correlated with a partition of N by the following map:

Definition 5.24 (Map ρ). Where KS is the set of known synonym sets, N is a set of co-referring name types and $\Phi(N)$ is a partition on N let ρ be a map

$$\rho : KS \longrightarrow \Phi(N)$$

such that $\rho(\bar{n}) = \{[\dot{n}] : [\dot{n}] \subseteq \bar{n}\}$. ◄

The map ρ assigns to each KS set $\bar{n}$ the set of name types the tokens of which are known by the agent to co-refer.

5.6.0.1 Partitions and Inferential Competence.

It can now be illustrated how each possible partition of N can be regarded as a way of being inferentially competent. The guiding idea is that every partition of $N = \{[\dot{n}_1], [\dot{n}_2], ..., [\dot{n}_k]\}$ can be seen as encoding information about the agent's knowledge: where two name types are in different cells in the partition, this reflects that the agent does not know that the tokens of them co-refer (which, in fact, they do as all the name types in N are synonyms). If, on the other hand, two name types are in the same cell of the partition, the agent knows the tokens of these types to co-refer. Hence, a finer partition of N means less knowledge; a coarser partition means more knowledge. This is exemplified thrice:

First, regard a situation in which agent i has only the inferential knowledge assigned by default. That is, assume that for all $n \in LEX$, the agent is only *trivially competent.* Then this way of being inferentially competent corresponds to the partition of N into singletons: each cell consist of exactly one name type, the tokens of which are known to co-refer (by virtue of syntactic competence). More specifically, the connection between trivial competence and the singleton partition of N can be presented as follows:

If i is trivially competent with all n, then for all $n \in [\dot{n}_i]$, for all $[\dot{n}_i] \in N$, it will be the case that $\bar{n} = [\dot{n}_i]$ if $\mathcal{I}(n) = \dot{n}_i$. This follows since trivial competence, $M, w \models_v K_i(\mu(n) = \mu(n))$, implies knowledge of co-reference, $M, w \models_v K_i(\mu(n) = \mu(n'))$, for all names of the same type, i.e. if also $M, w \models_v (n = n')$. Put differently, trivial competence is equivalent to requiring that $M, w \models_v K_i(\mu(n) = \mu(n'))$ iff $M, w \models_v (n = n')$. Hence $n' \in \bar{n}$ iff $\mathcal{I}(n') = \mathcal{I}(n)$. Therefore, if $\mathcal{I}(n) = \dot{n}_i$, then $\mathcal{I}(n') = \dot{n}_i$, so by Definition 5.21 $n' \in [\dot{n}_i]$. So $\bar{n} = [\dot{n}_i]$.

Now regard the map ρ. For the requirement to be satisfied in the present case, the partition $\Phi(N)$ has to be one where each cell is a singleton, and $\rho(\bar{n}) = \{\bar{n}\} = \{[\dot{n}]\}$. So, to summarize, where KS is given by an agent that is only trivially competent, and this set is related to a partition $\Phi(N)$ of N, this partition must consist of singletons, i.e. $\Phi(N) = \{\{[\dot{n}_1]\}, \{[\dot{n}_2]\}, ..., \{[\dot{n}_k]\}\}$.

Second, assume that i knows of two synonyms that they co-refer, but is only trivially competent with the remaining names. That is, assume that i is trivially competent with all names like above, except for the names from exactly two different types, which the agent knows to co-refer. That is, for names $n \in [\dot{n}_j]$ and $n' \in [\dot{n}'_l]$ such that $[\dot{n}_j] \neq [\dot{n}'_l]$, it is the case that $M, w \models_v K_i(\mu(n) = \mu(n'))$. In this case, the intuition is that the agent knows more, so this should correspond to a coarser partition of N. This is indeed the case:

Under the assumption that $M, w \models_v K_i(\mu(n) = \mu(n'))$, $\bar{n} = \bar{n}'$. Assume that $n \in [\dot{n}_1]$ and $n' \in [\dot{n}_2]$. Then $[\dot{n}_1] \subseteq \bar{n}$ and $[\dot{n}_2] \subseteq \bar{n}'$. Therefore, $\rho(\bar{n}) = \{[\dot{n}_1], [\dot{n}_2]\}$. The cell does not contain further elements as no further n'' from different name types are known synonyms to either n or n'. Further, as i was assumed trivially competent with such remaining n'', where $n'' \in [\dot{n}_i]$, $\rho(\bar{n}'') = \{[\dot{n}_i]\}$. Hence, for ρ to fulfill the requirement in the definition, the partition of N must be $\Phi(N) = \{\{[\dot{n}_1], [\dot{n}_2]\}, \{[\dot{n}_3]\}, ..., \{[\dot{n}_k]\}\}$.

Third, assume i knows that all type names in N co-refer. This should correspond to the coarsest possible partition, the trivial partition $\Phi(N) = \{N\}$. This is the case: from the assumption that i knows that all names used in N co-refer, it follows that $\bar{n} = \bar{n}'$ for all such n, n'. Hence $\rho(\bar{n}) = \rho(\bar{n}')$ for all such n, n'. For ρ to map all $\bar{n}$ to the same partition cell while maintaining the defining requirement, this partition cell must contain all $[\dot{n}_i]$. The only partition containing this cell is the trivial partition, $\Phi(N) = \{N\}$.

5.6.0.2 Size of the Hierarchy.

The number of possible ways to be inferentially competent with respect to a set of synonyms N with cardinality $|N| = n$ is equal to the sum of the sets *Sterling numbers of second kind.* The Sterling number of second kind with parameters n and k is the number of possible partitions with k cells of a set with n elements, cf. (Forst, 2006, sec. 9.1). This number is denoted $\{^n_k\}$ and the number of possible partitions is thus

$$\sum_{k=1}^{n} \left\{ \begin{array}{c} n \\ k \end{array} \right\}$$

These numbers get quite large, quite fast. To illustrate, for a set consisting of 3 synonyms, there are

$$\sum_{k=1}^{3} \left\{ \begin{array}{c} 3 \\ k \end{array} \right\} = \left\{ \begin{array}{c} 3 \\ 1 \end{array} \right\} + \left\{ \begin{array}{c} 3 \\ 2 \end{array} \right\} + \left\{ \begin{array}{c} 3 \\ 3 \end{array} \right\} = 1 + 3 + 1 = 5$$

possible partitions. Where there are 5 co-referring names, this number grows to 52, and where 10 synonyms exist, there are 115975 different ways to be inferentially competent.[10] This number seems too large for the hierarchy to be applicable in practice, in particular since these are found without taking application and correlation into consideration.

5.6.0.3 Strengths.

The idea that finer/coarser partitions of N encode that the agent has less/more information about the co-reference of synonyms implies two things: first, that a partial order of different strengths can be found and second, that some of these strengths may be equal.

Regard for example a set of synonyms $N = \{[\dot{n}_1], [\dot{n}_2], [\dot{n}_3]\}$. Then the two partitions $\Phi_{2+1}(N)$ and $\Phi_{1+2}(N)$

$$\begin{array}{ll} \Phi_{2+1}(N) = & \{\{[\dot{n}_1], [\dot{n}_2]\}, \{[\dot{n}_3]\}\} \\ \Phi_{1+2}(N) = & \{\{[\dot{n}_1]\}, \{[\dot{n}_2], [\dot{n}_3]\}\} \end{array}$$

can be regarded as equal in strength, since in both cases, the agent knows that two out of three synonyms co-refer. The partitions

$$\Phi_3(N) = \{\{[\dot{n}_1], [\dot{n}_2], [\dot{n}_3]\}\}$$

and

$$\Phi_{1+1+1}(N) = \{\{[\dot{n}_1]\}, \{[\dot{n}_2]\}, \{[\dot{n}_3]\}\}$$

are respectively stronger and weaker, as the agent knows that respectively more and less names are synonymous.

---

[10]A recursive formula for calculating Sterling number of second kind can be found in (Forst, 2006). An explicit formula can be found online on Wikipedia.

For a set of co-referring names N with cardinality $|N| = k$, the number of different strengths is given by the *partition number* $p(k)$ of k, i.e. the number of ways k can be represented as a sum of natural numbers where the order of the summands is irrelevant, see (Forst, 2006, sec. 9.6). This was illustrated by the subscripts above: $\Phi_{2+1}(N)$ is the partition of the three object set N into two sets with cardinality 2 and 1. This is of same strength as the partition $\Phi_{1+2}(N)$, and if the order of the summands in the subscripts are irrelevant, these two partitions will count only once.

For the values 3, 5 and 10, the partition numbers are, respectively, $p(3) = 3$, $p(5) = 7$ and $p(10) = 42$. Hence, focusing on the various strengths of inferential competence greatly diminishes the size of a possible hierarchy.

5.6.0.4 Adding Application.

The basic idea for adding application to the picture is that where an agent is equally inferentially competent, but is able to apply more names in one case, this case will be stronger in strength. Where the agent is more inferentially competent in one case, but unable to apply names, and less inferentially competent in another, where the agent is able to apply a name, these will be considered incomparable. See Figure 5.6.1 for an illustration.

When adding application, the combinatorial task of finding the cardinality of the partially ordered hierarchy becomes difficult. By Fact 5.1 and Fact 5.2, for any inferential competence strength, the agents can at most apply names from one of the cells without part of the partition collapsing. For example, if an agent can apply both $n \in [\dot{n}_1]$ and $n' \in [\dot{n}_3]$, but is supposed inferentially competent of strength $\Phi_{2+1}(N)$ above, an inconsistency can be reached. By fact Fact 5.1, the agent must be inferentially competent with the strength of $\Phi_3(N)$. This insight is important, as it eliminates many superfluous 'strengths'.

It is conjectured that the cardinality of the hierarchy of application and inferential competence can be found in the following way: for a set of co-referring names N, identify each of the possible partitions of $|N|$, where the order of the summands are irrelevant. For each of these partitions, identify the number of *different* summands strictly greater than 0 and add 1. Finally, sum all these numbers. Exactly *how* this is to be done if not manually is unknown at this juncture.

The method will correctly describe the different types of strengths since each partition with n different summands gives rise to $n + 1$ different strengths: one for each summand a name of which the agent can apply, plus one where the agent can apply no names. The joint sum will hence add to the different strengths possible.

The hierarchy for $|N| = 3$ is illustrated in Figure 5.6.1, and for this case, the method outlined is applied as follows. First, there are three different ways to partition 3: 3, $2 + 1$ and $1 + 1 + 1$. The number of different summands plus 1 is then 2, 3 and 2. The sum is 7, which (at least in this case) coincides with the amount of different elements in the hierarchy.
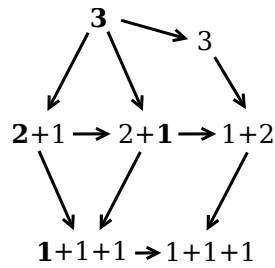
$$
\begin{array}{ccc}
 & \mathbf{3} \longrightarrow 3 & \\
 \swarrow & \downarrow & \searrow \\
\mathbf{2}{+}1 \longrightarrow 2{+}\mathbf{1} \longrightarrow 1{+}2 & & \\
\searrow & \downarrow & \swarrow \\
 & \mathbf{1}{+}1{+}1 \longrightarrow 1{+}1{+}1 &
\end{array}
$$

Figure 5.6.1: Hierarchy for application and inference, for $|N| = 3$. Numbers in bold mark partition cells for which the agent can apply a name, arrows indicate strength order. Notice $2{+}\mathbf{1}$ and 3 are not connected, illustrating that the strength order is partial.

Conclusion. The partial order illustrated in Figure 5.6.1 seems to suggest that though it may become very large in many applications, a hierarchy of strengths composed of inferential competence and application does seem possible to construct. Given a suitable notation, it further seems that the combinatorial aspects of the order are solvable, provided that the loose method provided above is feasible. Further, given such a suitable notation, a taxonomy of semantic competence strengths seems constructible. Further, Figure 5.6.1 seems to illustrate the possibility of constructing a lexicographic ordering of such competence strengths.

Further aspects should be included, though, if the hierarchy are to elucidate *all* possible aspects of semantic competence. The inclusion of these are beyond the scope of the present appendix. That a tentative order of semantic competence type and strengths can be constructed provides credence to the hypothesis that an appropriate hierarchy and taxonomy can be constructed in general.

6 Showing Proof of Concept

In section 2.1.2 the dilemma put to the Millian theory of meaning by Frege was introduced (restated below). Being a dilemma, it presented two options for the Millian to interpret identity statements. It is usually argued that neither of these are feasible, and that Millianism should therefore be rejected. In the present chapter, each of the two disjuncts will be regarded in the light of the theory of semantic competence developed in the previous chapter. It is argued in each case that the more fine-grained view on informational content the theory of semantic competence applies yields analyses of the arguments which elucidate the informational structure. The revealed structure results in a view of the two cases as transparent, non-problematic linguistic situations.

6.1 Frege's Puzzle about Identity

As mentioned in section 2.1.2, the *Puzzle about Identity* was raised against Millianism based on the interpretation of the identity expressed in the identity statements as being a relation between *objects*. In the mentioned section, using the identity statements

(a) Hesperus is Hesperus

(b) Hesperus is Phosporus

the argument was presented in the following way:

(A) (a) and (b) mean the same.

(A→B) If (a) and (b) mean the same, then a semantically competent speaker would know that (a) and (b) mean the same.

(B→C) If a semantically competent speaker would know that (a) and (b) mean the same, then they are equally informative to the speaker.

(¬C) (a) and (b) differ in informativeness to the competent speaker.

∴ Contradiction.

The four premises are jointly inconsistent, and the typical textbook choice is to reject premise (A). This premise is a consequence of the Millian view, and the conclusion drawn is hence that there must be more to meaning than mere reference.

Given the formal framework from the preceding chapter, the argument can now be given a rigorous analysis.

Assume that the identity statements 'Hesperus is Hesperus' and 'Hesperus is Phosphorus' are captured by $(\mu(n) = \mu(n))$ and $(\mu(n) = \mu(n'))$, respectfully. The premise (A) states that these have the same meaning. By truth-theoretic semantics, this means that they should have the same truth conditions, so assume

$$(\mu(n) = \mu(n)) \leftrightarrow (\mu(n) = \mu(n')) \tag{6.1.1}$$

holds in the actual world w of a $\mathcal{C}_{\mathsf{2QEL}}$ model, M. As the left-hand side is a validity, this is equivalent to the substantial assumption that

$$M, w \models_v (\mu(n) = \mu(n')) \tag{6.1.2}$$

The second premise is that (6.1.2) implies that any competent speaker knows (6.1.1).

The truth of this premise depends on the type of competence meant. The last three premises of the argument will be run through using *inferential competence*, *application* and *correlation*. The ability to name objects is not relevant, as this ability assumes agents being presented with objects rather than names.

Inferential Competence   The second premise states that if n and n' mean the same, i.e. $(\mu(n) = \mu(n'))$, and agent i is inferential competent with respect to the two names, then agent i would know that n and n' mean the same. As i is inferentially competent with respect to n iff

$$M, w \models_v (\mu(n) = \mu(n')) \rightarrow K_i(\mu(n) = \mu(n'))$$

for all n', cf. (7.3.2) on page 120, the premise holds true, so

$$M, w \models_v K_i(\mu(n) = \mu(n')) \tag{6.1.3}$$

The third premise states that (6.1.3) implies that the two identity statements are equally informative to the agent. 'Equally informative' is taken to mean that the two statements would eliminate the same worlds from agent i's model *if truthfully announced to the agent*, in the sense of (van Ditmarsch et al., 2008).

As $(\mu(n) = \mu(n))$ is a validity, it eliminates no worlds, so the premise can be reduced to (6.1.3) implying that

$$\neg \exists w' \sim_i w : M, w' \models_v \neg(\mu(n) = \mu(n')) \tag{6.1.4}$$

That no $\neg(\mu(n) = \mu(n'))$ world exists follows from (6.1.3) and the semantics of the K_i operator. Hence the third premise holds true as well for models in which the agent is assumed to be inferentially competent.

This is not the case with the last premise, namely that the identity statements should *not* be equally informative, i.e.

$$\exists w' \sim_i w : M, w' \models_v \neg(\mu(n) = \mu(n')) \tag{6.1.5}$$

This premise is false as a direct consequence of the assumption of inferential competence.

To evaluate, if one sticks with Millian meaning and assume agent i inferentially competent, i does not learn anything new by being told that the two names co-refer. This is a clear consequence of the information the agent is assumed to possess in virtue of being inferentially competent. Hence, the conclusion that the agent is not informed by the identity statement seems far less 'puzzling' due to the transparency of the competence type when formalized. Further, it does not seem paradoxical enough (if at all) to warrant a rejection of the Millian view.

Referential competence, application   The last three premises are assessed under the assumption that agent i is able to apply both n and n', i.e. that Definition 5.15 on page 91 holds at w for both.

The second premise amounts to

$$(\mu(n) = \mu(n')) \wedge \exists x K_i(\mu(n) = x) \wedge \exists y K_i(\mu(n') = y) \rightarrow K_i(\mu(n) = \mu(n'))$$

which is valid on the class of models, as was stated as Fact 5.1 on page 97. Hence, from the assumption of the antecedent, it follows that i is inferentially competent with respect to the two names, i.e. (6.1.3) holds. As above, this entails the truth of (6.1.4) and the falsity of (6.1.5). Hence, the agent *will not* be informed by the identity statement.

This is an unsurprising conclusion if one notices the informational structure behind it. For the ability to apply a name amounts to the agent being able to *identify the referent*, in the sense of having appropriate *de re* knowledge. Under suitable assumptions regarding short-term memory, a statement that two names co-refer will clearly not be informative when the agent has just identified the referent of each name.

In conclusion, if Millianism is assumed in conjunction with the agent being able to apply both names, the agent is not informed by the identity statement. Again, the formal analysis shows that there is nothing puzzling or paradoxical about the agent not being informed, and this version of Frege's Puzzle about Identity does not provide a strong argument against Millianism.

Referential competence, correlation   Running through the argument using the weaker ability of correlation, the second premise becomes

$$K_i(\mu(n) = a) \wedge K_i(\mu(n') = b) \rightarrow K_i(\mu(n) = \mu(n')) \tag{6.1.6}$$

This formula is satisfiable, but not valid in the class of $\mathcal{C}_{2\text{QEL}}$ models. This implies that (6.1.3) will be true or false depending on the specific model. In case the

consequent of 6.1.6 is satisfied, the agent will have knowledge of co-reference, and it will, like above, not be surprising that he is not informed by the identity statement. In case the consequent fails, a new situation arises. In particular, this will imply that (6.1.4) likewise fails to be the true. From this it follows that the 'intuitive premise'

$$\exists w' \sim_i w : M, w' \models_v \neg(\mu(n) = \mu(n'))$$

is now satisfied, as opposed to the above cases (the above formula is (6.1.5) from above). This in turn means that the agent may in fact be informed by the identity statement. If a truthful announcement of the identity statement was made to the agent, any w' as specified in (6.1.5) could be eliminated, and the agent would thereby gain information.

By the truthful announcement, the agent is informed on both an *inferential* and a *conceptual* level. First, the agent will after the announcement have knowledge of co-reference with respect to the two names. Second, where the agent before had two distinct concepts, the agent's concepts of a and b will after the announcement have merged.

However, given the weaker notion of competence, this does not conflict with the assumption of Millian meaning of proper names. To see this, notice that the two premises (A→B) and (B→C) from the argument above are false when assuming the weaker semantic competence type: correlation. Hence, the original contradiction can no longer be derived. Therefore, again, this version of the argument does not warrant a rejection of Millianism.

In conclusion   Neither of the above three versions of Frege's Puzzle about Identity provided a proper argument against Millianism. In fact, when the structure of the situations where analyzed using precise notions of semantic competence and clear assumptions regarding the agent's information, all three situations seemed natural: where the agent was not informed, there was a clear, epistemic reason, and where the agent was informed i) the specific way he was informed could be spelled out, and ii) the case caused the original argument against Millianism to break down. In short, neither of the three types of competence result in a strong argument against the Millian view, and the utilization of a formal theory of semantic competence makes this plainly visible.

6.1.1 Objection

Against the proposed analysis and resolve of Frege's Puzzle, one may object that the premises of the model misrepresent a very relevant case.[1] To exemplify, consider an astronomer originally informed by 'Hesperus is Phosphorus'. It is conceivable that this agent was in fact able to identify both Hesperus and Phosphorus, but still unaware that Venus is the referent of both names. Hence, the agent should be able

---

[1]This is not an objection from the general literature, but a concern of the author.

to apply both names, but still be informed by the identity statement. In short, one may object to the intuitive validity of

$$(\mu(n) = \mu(n')) \wedge \exists x K_i(\mu(n) = x) \wedge \exists y K_i(\mu(n') = y) \to K_i(\mu(n) = \mu(n')).$$

The example motivating this objection features an important aspect not explicitly mentioned, which is the *context of identification*. The cases reviewed in (Marconi, 1997) relate only to a single such context: the experiments where not conducted in a variety of different contexts. The astronomer example, on the other hand, features multiple contexts. In particular, astronomers where able to identify Hesperus in the evening, whereas Phosphorus was identified in the morning, and the astronomers where unaware that the two objects identified where in fact the same (i.e. in neither context where they able to identify Hesperus and Phosphorus as the same object).

In order to show that adding the extra parameter of *context* does not change the conclusions drawn in the above section, the following chapter will extend the present model for semantic competence to one where multiple contexts can be modeled.

However, before moving on to this extension, it will be argued that the second disjunct of Frege's Dilemma also benefits from an analysis in terms of semantic competence and informational content.

6.2 The Problem of Non-Informativeness

As mentioned in section 2.1.2, the second disjunct of Frege's Dilemma concerns the non-informativeness of an identity statement where the informational content of 'n_1 is n_2' is taken to be that the two names co-refer. The critique raised was that the informational content will then be vacuous, in that it will regard linguistic convention rather than word-world relations (unless it is already known which object one of the names refer to), and that this disjunct therefore was unfeasible.

In the above, it has silently been assumed that the way the identity statements where modeled corresponded with the assumptions of the first disjunct, namely utilizing the identity as a relation between objects. This is indeed also the case, as $\mu(n_i) \in Obj$. Yet, the same modeling means that the second disjunct of Frege's Dilemma may also be fulfilled. This is so as the identity statement $\mu(n_1) = \mu(n_2)$ may be read 'the meaning of the name n_1 is the same as the meaning of the name n_2'. Hence, the statement expresses that

> the sign '$[n_1]$' designates the same object as the sign '$[n_2]$' (Collin and Guldmann, 2010, p. 49-50).

This means that the second disjunct is in effect – at least partly and then unproblematically, it will be argued, as this *depends on the semantic competence of the speaker*.

In order to analyze this second problem, it cannot be assumed that the agent can apply either name, and neither can it be assumed that the type of competence

meant is inferential competence. It cannot be the case that the agent can apply one name as this is precluded by the argument. Indeed, if the agent was able to apply one name and subsequently be informed by the identity statement, then this would result in knowledge of a word-world relation, as the agent would come to be able to apply the second name, cf. Fact 5.2 on page 97. If the agent was already inferentially competent with respect to the two names, the announcement would be unsurprisingly uninformative, cf. the previous section. Hence, the agent is weaker semantically competent, i.e. has some sort of correlational competence.

The case where the agent non-trivially correlates n_1 and n_2 with ambiguous concepts a and b was discussed above, and it was seen that the agent was informed *both* about linguistic convention, but also conceptually. Hence, in that situation, the announcement of the identity statement *did not* have a trivial effect, contrary to the objectionable conclusion of the Problem of Non-Informativeness.

The weaker choices left are that the agent is *trivially competent*, cf. (5.5.4) on page 96, with respect to either one or both names.

Trivially competent with one name   In the case where i correlates n_1 with a and is trivially competent with respect to n_2, where $M \in \mathcal{C}_{\mathsf{2QEL}}$ with actual world w, it will be the case that

$$M, w \models_v K_i(\mu(n_1) = a) \land K_i(\mu(n_2) = \mu(n_2)).$$

When the announcement is made of the identity statement, the agent will again gain both conceptual and linguistic information. The conceptual information will result in the knowledge that n_2 is another name for a, i.e.

$$M, w \models_v K_i(\mu(n_2) = a)$$

Hence, though $C_i^{a,w}(\mathcal{I}(a, w)) > 1$ by assumption, i.e. i has an ambiguous concept for a at w, the announcement will still result in a new relation between the word lexicon, the semantic lexicon and real-world objects. Thus the identity statement *does not* lack proper informational content when this type of semantic competence is considered. Again, the objectionable conclusion is not reached.

Trivially competent with both names   In order to construct a proper argument from the second disjunct, it must therefore be assumed that the agent is trivially competent with respect to *both* names. This amounts to regarding *any* $\mathcal{C}_{\mathsf{2QEL}}$ model. Every world in such will satisfy

$$M, w \models_v K_i(\mu(n_1) = \mu(n_1)) \land K_i(\mu(n_2) = \mu(n_2)).$$

Hence the notion of *trivial* competence. Assume that $(\mu(n_1) = \mu(n_2))$ is satisfied at w in order for the possibility of a truthful announcement of it. After the announcement, i will come to know that the two names co-refer, i.e. will gain information about linguistic convention in the sense of being inferentially competent with the

two names, but will still not be able to apply either name. Hence, a new relation is *not* drawn from the word lexicon, through the semantic lexicon and onto a real-world object. This is exactly what the problem on non-informativeness prescribes and raises as a complaint against the used interpretation of identity statements.

But that the agent only gains abstract, linguistic knowledge is unsurprising when a precise version of the premises of the argument is taken into consideration. When the agent is assumed to be *completely* uninformed, the update has *no* other information to interact with, and the conclusions learned by the agent are hence limited – in this case, to some concerning only linguistic phenomena.

This does not seem to pose a problem once properly understood, though. In fact, if the Millian theory of meaning entailed that an agent utterly uninformed of the meaning of two terms, by being told that they co-referred where then able to *identify the referent by default*, this would seem to be a serious problem. Specifically, this would mean that the theory entailed that inferential competence with two names implies the ability to apply both names, which is inconsistent with the findings reviewed in (Marconi, 1997).

Thus, when the information possessed by the agent in the announcement situation is modeled explicitly, it is seen that it is both unsurprising and unproblematic that, until the agent is able to relate either name to the world, the agent will be in possession

> of an item of information about two languages and not knowledge about the world. (Collin and Guldmann, 2010, p. 50)

6.3 In Sum

In this chapter, the two disjuncts of Frege's Dilemma has been analyzed using the constructed formal theory of semantic competence. This has been done by analyzing both disjuncts using the different notions of semantic competence extracted from this theory. For each disjunct, and each type of semantic competence, it has been argued that the case did not provide an argument against Millianism. Ignoring the objection of section 6.1.1, one can draw the overall conclusion that Frege's Dilemma *does not* pose a problem for the Millian theory of meaning.

This conclusion stands in sharp contrast with the standard textbook conclusion. The reason for the different conclusion drawn here follows from the application of a strictly defined theory of semantic competence. The theory provides precise notions of semantic competence based on the agent's epistemic situation. This differs from the textbook approach, where the notion of semantic competence is left undefined. Further, the notion of *being informed* is clear in relation to the formal theory, which results in a transparency of 'intuitive' textbook premises. This transparency in makes these premises less intuitively correct. As a result, the formal theory allows for more detailed analyses of the disjuncts, showing that the apparent inconsistency with Millianism is non-existent. In sum, it has been shown that the situations described as puzzling and problematic in the literature are in fact natural

and well-understood if the agents information regarding the used language is taken into consideration.

7 Contextual Semantic Competence

In section 6.1.1, an objection was raised against the analysis of Frege's *Puzzle about Identity*. The objection was in relation to the analysis in terms of *application*. It was based on the intuitive idea that though an agent is able to identify the referent of two co-referring names, the agent might not know that these two referred to the same object. As a reply it was argued that this presumed that the agent where doing the identification in *distinct contexts*, and that once such was added to the formal model, the analysis would still be applicable. To lend credence to this argument, this chapter adds contexts to the $\mathcal{C}_{2\mathsf{QEL}}$ models. The formal details are introduced in section 7.1. The constructed model class is discussed generally in section 7.2, and specifically in relation to semantic competence in section 7.3. Finally, the objection against the previous analysis is review in section 7.4.

A thorough discussion of the nature of contexts will not be presented. It is assumed that a context is an abstraction of a class of situations, prominent features of which reoccur enough for such to be made. The actual world of a context is then understood as the world encoding what is actually the case in these situations. It is assumed that the actual worlds are metaphysical possibilities relative to one another, where these worlds consist of the same elements, which persist across worlds.

Examples of contexts may for instance be '*at work*', '*at home*' and '*on weekend trip*', or '*stargazing in the morning*' and '*stargazing in the evening*'.

7.1 Adding Contexts

The $\mathcal{C}_{2\mathsf{QEL}}$ models are single context models. They contain *one* actual world and epistemic alternatives to this world. Adding further context requires adding further actual worlds, each capturing what is the actual state of affairs in the appropriate context. Further, for each such added actual world, the model must contain a set of epistemic alternatives. If it is assumed that the agents should be able to tell which context they are in, the epistemic alternatives for a context should not overlap with such from any other. This assumption will be made here.

Definition 7.1 (**Context Structure**)**.** Where W is a non-empty set of worlds, a *context structure on W* is a pair (S, Act) where S is a partition of W

$$S = \{S_1, S_2, ..., S_n\}$$

where each S_k contains an actual world ω_k from the set of *actual worlds*,

$$Act = \{\omega_1, \omega_2, ..., \omega_n\}.$$

Each $S_k \in S$ will be referred to as a *context* from (S, Act). ◀

A context structure defines which worlds from W are of the same 'context type' as each actual world. It will be assumed that agents are capable of telling which context they are in, and it will therefore be required that the indistinguishability relations must induce an at least as fine partition on W:

Definition 7.2 (**Context Distinguishability**)**.** Where S_k is a context from (S, Act), an indistinguishability relation $\sim_i$ that *distinguishes contexts* satisfy

$$\text{If } w \sim_i w' \text{then } w, w' \in S_k.$$

Define the set of agent *i's epistemic alternatives to* ω_k by

$$S_k^i = \{w : (\omega_k, w) \in \sim_i\}$$

◀

That is, where w and w' are indistinguishable to agent i, these are instances of the same context. In other words, the agents' epistemic alternatives are pre-partitioned by the modeler in order to ensure that the agents know which context they are in.

That contexts are disjunct has the effect that the agents have no conception of what is the case in other contexts. In fact, as such contexts are completely disconnected, nothing in one context will have any bearing on the truth of formulas in other contexts. To add a connection between contexts, two alethic-type relations between actual worlds are defined:

Definition 7.3 (**Objective Possibility**)**.** Where Act is the set of actual worlds from (S, Act), define the *objective possibility relation* by

$$\mathcal{R} = Act \times Act.$$

◀

The objective possibility relation relates all and only the actual worlds. A world accessible by this relation is interpreted as *the* way the world actually is, when in a different context. As such, the relation is not between metaphysical *alternatives*, but between metaphysical *actualities*.[1]

The objective possibility relation hence ranges over *actual* worlds, but since an agent may be in doubt as to which world is the actual in some context, the agent's *conception* of such possibilities also needs to be defined:

---

[1]Recall that the contexts are assumed to be *abstractions* of sets of 'real' contexts, including only re-occurring features. Hopefully the reader agrees that problems regarding spatio-temporal co-existence and specific metaphysical nature of such 'co-existing' possible worlds can be left undiscussed.
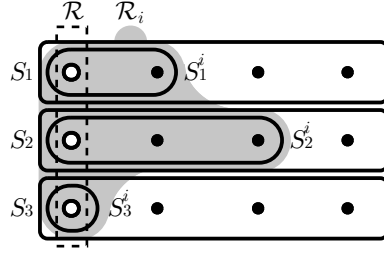
Figure 7.1.1: A $\mathcal{C}_{2\mathsf{QEL}}^{\square}$ model with three contexts, each containing i) one actual world (white center), ii) three epistemic alternatives and iii) one information cell for agent i. The objective possibility relation $\mathcal{R}$ is marked by the dotted line, and the subjective possibility relation $\mathcal{R}_i$ by the gray area.

Definition 7.4 (Subjective Possibility). Where $S = \{S_1, ..., S_n\}$ is the partition from (S, Act), the *subjective possibility relation for agent i* is defined by

$$\mathcal{R}_i = \bigcup_{k \leq n} S_k^i \times \bigcup_{k \leq n} S_k^i.$$

◀

The definition states that any epistemic alternative to any actual world is considered a possibility from any other epistemic alternative.

The subjective possibility relation is used to express the agents' conception of the objective possibility relation. The idea behind the definition is that where an agent cannot tell the worlds in S_j^i from ω_j and cannot tell the worlds in S_k^i from ω_k, and ω_j and ω_k are possible from one another, then all the worlds in S_j^i should be considered possible from all the worlds in S_k^i and *vice versa*.

Definition 7.5 (Context Model). A *context model M* is a tuple

$$M = \langle W, (S, Act), (\sim_i, \mathcal{R}_i)_{i \in I}, \mathcal{R}, Dom, \mathcal{I} \rangle$$

where W is a set of states, (S, Act) is a context structure on W, $(\sim_i, \mathcal{R}_i)_{i \in I}$ is a set of indistinguishability relations distinguishing contexts and subjective possibility relations, one of each for each $i \in I$, $\mathcal{R}$ is an objective possibility relation over Act, and Dom and $\mathcal{I}$ are as in $\mathcal{C}_{2\mathsf{QEL}}$ models.

The class of context models is denoted $\mathcal{C}_{2\mathsf{QEL}}^{\square}$. ◀

7.1.1 Syntax and Semantics

The language $\mathcal{L}_{2\mathsf{QEL}}$ is augmented with two operators $\square$ and $\square_i$, $i \in I$. The resulting language is denoted $\mathcal{L}_{2\mathsf{QEL}}^{\square}$. The set of $\mathcal{L}_{2\mathsf{QEL}}^{\square}$ well-formed formulas include the well-formed formulas of $\mathcal{L}_{2\mathsf{QEL}}$, and where φ is a well-formed formula, so is

$$\square\varphi \mid \square_i\varphi$$

Dual operators of $\Box$ and $\Box_i$ are defined by $\Diamond := \neg\Box\neg$ and $\Diamond_i := \neg\Box_i\neg$.

The semantics for connectives, quantifiers and the knowledge operators remain as defined in chapter 3 and section 5.2. The semantics for the necessity operators are defined as follows:

$$M, w \models_v \Box\varphi \quad \text{iff} \quad \forall\omega \in Act, \, M, \omega \models_v \varphi$$
$$M, w \models_v \Box_i\varphi \quad \text{iff} \quad \forall w' : w\mathcal{R}_i w' \Rightarrow M, w' \models_v \varphi$$

The reading of the box operators are 'in all contexts, φ' and 'in all contexts, for all i knows, φ', respectfully, and the diamond operators are read 'in at least one context, φ' and 'in at least one context, for all i knows, φ'. Having the comments regarding the loose formulation of the involved metaphysics in mind, the readings can be put as 'objectively/subjectively necessarily/possibly, φ', respectively.

7.1.2 Restrictions on Meaning and Identity

As mentioned in section 5.2, the meaning function utilized there made names rigid in the sense that they had the same meaning in all metaphysically possible worlds. Of such metaphysically possible worlds, there was possibly only one, namely the actual world. As $\mathcal{C}_{\text{2QEL}}^{\Box}$ include more than one actual world, and hence more than one metaphysically possible world, the meaning function must be restricted to preserve rigidity. To this end, it is required of μ that

$$\forall\omega, \omega' \in Act : M, \omega \models_v (\mu(n) = a) \Rightarrow M, \omega' \models_v (\mu(n) = a) \tag{7.1.1}$$

for any $a \in OBJ$ and all $n \in LEX$. The requirement states that $\mu(n)$ must be the same in all actual worlds for all names n, i.e. all names refer rigidly over Act. Adopting this restriction results in the local validity over all $\omega \in Act$ of

$$(\mu(n) = \mu(n')) \rightarrow \Box(\mu(n) = \mu(n'))$$

i.e. that co-reference of names is objectively necessary.

The currently defined semantics does not distinguish between object identity across actual worlds and epistemic alternatives, though one rests on physical duration and the other on current information. One cross-identity restriction regarding physical identity statements, namely their persistence across actual worlds, will be assumed:

$$\forall\omega, \omega' \in Act : M, \omega \models_v (a = b) \Rightarrow M, \omega' \models_v (a = b) \tag{7.1.2}$$

This requirement ensures that objects are self-identical across actual worlds. Hence, across actual worlds, an incontingent identity system is in effect. Given (7.1.2), *necessity of identity*

$$(t = t') \rightarrow \Box(t = t') \tag{7.1.3}$$

is satisfied at all $\omega \in Act$ for all terms in which function symbols do not occur. For $w \notin Act$, (7.1.3), may not hold, but the weaker

$$\Diamond(t = t') \rightarrow \Box(t = t') \tag{7.1.4}$$

is valid for all terms t, t' in the class $\mathcal{C}^\square_{\mathsf{2QEL}}$ under the assumptions listed above.

In the following, if not otherwise specified, reference to the class $\mathcal{C}^\square_{\mathsf{2QEL}}$ is meant to include the above restrictions. Proofs of validity and invalidity of mentioned formulas will not be shown, due to considerations of space.

7.2 Axioms and Operator Interplay

The following section has two purposes. The first is to illustrate the features of the theory of contextual semantic competence that can be constructed on the basis of the $\mathcal{C}^\square_{\mathsf{2QEL}}$ models. The second is to investigate the possibilities for constructing a complete axiom system for such a theory.

In order to find a complete axiom system, axioms for the separate operators must be present. With respect to the K_i- and $\square_i$-operators, then these validate the S5 axioms. However, this is not the case with the $\square$-operator. In particular, T does not hold for $\square$ as $\mathcal{R}$ is not defined to be reflexive in $W \setminus Act$. $\square$ validates the KD45 axioms[2], and in all $\omega \in Act$, T holds. It is therefore conjectured that a complete axiom system must include the S5 axioms for K_i and the KD45 axioms for $\square$. However, the S5 axioms for $\square_i$ may not be required, due to the possibility of defining $\square_i$ in terms of K_i and $\square$, as will be discussed below.

The models of $\mathcal{C}^\square_{\mathsf{2QEL}}$ exhibit some properties regarding the interplay of the three operator types defined.[3] First, the system allows that necessary truths are unknown. This is embedded in the invalidity of the schema

$$\square\varphi \to K_i\varphi \tag{7.2.1}$$

The feature is wanted as necessary *a posteriori* propositions and agent's (lack of) knowledge of these will be modeled. Necessary *a posteriori* propositions can be unknown, and the validity of (7.2.1) would preclude this possibility. Related, the schema

$$\square\varphi \to \square_i\varphi \tag{7.2.2}$$

is also invalid. As the $\square_i$-operator is meant to capture agent i's subjective view of what is necessary in all contexts, the invalidity of (7.2.2) nicely captures that this notion is not tainted in too high a degree by what is actually the case.

It may further be noted that

$$\square_i\varphi \to \square\varphi \tag{7.2.3}$$

is valid, and reflects that the subjectively possible worlds is a superset of the objectively possible worlds. The validity of (7.2.3) follows as $\mathcal{R} \subseteq \mathcal{R}_i$, for any $i \in I$.

---

[2]The KD45 axioms are K, 4, 5 and D, where D is $\lozenge\top$. D is valid in $\mathcal{C}^\square_{\mathsf{2QEL}}$ as there is always at least one actual world, cf. the definition of context structures above.

[3]That the following properties hold will not be proven, but can be shown simply by constructing satisfying models and counterexamples for contingent formulas or by deriving contradictions from the assumption of the negation of a formula stated to be valid.

Regarding the contrapositive of (7.2.3), $\Diamond\varphi \to \Diamond_i\varphi$, gives a nice, intuitive reading: whatever is objectively possible is subjectively possible. That is, no agent will gain knowledge in a way that will eliminate actual worlds from the set considered possible. It is interesting to note that the converse of (7.2.3) does not hold, and hence neither does it's contrapositive:

$$\Diamond_i\varphi \to \Diamond\varphi$$

That is, though agent i may conceive it possible that φ, this does *not* imply that φ is true in any of the actual possibilities.

Having as wide a semantic scope as it does, subjective necessity also implies knowledge, i.e.

$$\Box_i\varphi \to K_i\varphi \qquad (7.2.4)$$

is valid. This follows from the inclusion of $\sim_i$ in $\mathcal{R}_i$. The contrapositive of (7.2.4), $P_i\varphi \to \Diamond_i\varphi$, again yields an intuitive reading: if agent i considers φ possible in the actual context, then it is considered possible when the space of all contexts is under consideration. Hence, if i cannot rule out φ in the present context, then i cannot rule out φ in all contexts.[4] The converse implication, $K_i\varphi \to \Box_i\varphi$, is not valid.

Expressions of the type $\Box_i\varphi$ require truth of φ in a very big part of the model (all worlds connected to any $\omega \in Act$, if the mono-agent case is regarded), as was the reason for the two validities immediately above. The strong requirements for it's truth also result in the validity of

$$\Box_i\varphi \to K_i\Box_i\varphi$$

and

$$\Box_i\varphi \to \Box_i K_i\varphi$$

In general, $\Box_i\varphi$ hence implies φ prefixed with any n-length operator-block consisting of $\Box_i$- and K_i-operators. This can also be seen from (7.2.4) and axiom 4 for K_i and $\Box_i$.[5]

The $\Box_i$-operator may be defined using the $\Box$- and K_i-operators, by utilizing the following two validities

$$\Box K_i\varphi \to \Box_i\varphi \qquad (7.2.5)$$

$$\Box_i\varphi \to \Box K_i\varphi \qquad (7.2.6)$$

These validities are interesting from a meta-theoretical point of view. They show that the $\Box_i$-operator can be introduced as the abbreviation $\Box_i := \Box K_i$, as will be commented on below.

One unfortunate relation holds between the $\Box$- and K_i-operators. This is the validity of

$$\Box\varphi \to K_i\Box\varphi \qquad (7.2.7)$$

---

[4]Mind the quantifier scope: it is not the case that if i cannot rules out φ in the actual situation then in all situations, i cannot rules out φ. Though $\neg K_i\varphi$ holds in one partition, this does not preclude that $K_i\varphi$ holds in others.

[5]If 4 for $\Box_i$ is included.

As T is not valid for $\Box$, this does not imply that necessities are known, as mentioned above. Yet, the validity is counter-intuitive and readings its contrapositive does not clarify matters. The definition of the objective necessity operator hence results in an unwanted "tap" into knowledge of objective necessity. The best response available at the present is that formulas like (7.2.7) where $\Box$-operators occur within the scope of knowledge operators have *no coherent reading*, and that knowledge of necessities should be expressed using the for that purpose designed subjective necessity operators. It is noted that (7.2.7) is problematic, but this will be ignored as it will have no bearing on the following analysis.

Completeness

With the Canonical Class Theorem and Theorem 5.1, showing the completeness of $\mathbf{S5}_{n,\overline{\sigma}}$ with $\overline{\sigma} = \{\sigma_{OBJ}, \sigma_{LEX}\}$ with respect to $\mathcal{C}_{\mathsf{2QEL}}$, one could hope finding a complete axiom system for $\mathcal{C}_{\mathsf{2QEL}}^{\Box}$ was easy. Further, in order to obtain a logical theory of semantic competence with contexts, axioms characterizing the restrictions on meaning and identity should be found.

Regarding the latter, one could hope that adding 7.1.4 would be sufficient, as this seemingly enforces the requirements across *Act* by securing that meaning is rigid and identity constant across all actual worlds.

Certain aspects of the structure between $\sim_i$, $\mathcal{R}_i$ and $\mathcal{R}$ is seemingly captured by formulas in the section above: including 7.2.5 and 7.2.6 ensures that both $\sim_i$ and $\mathcal{R}$ are sub-relations of $\mathcal{R}_i$, and further allows for the definition of $\Box_i$ in terms of K_i and $\Box$. This means that extra axioms for the new modalities can be restricted to axioms for $\Box$.

Finding the appropriate axioms for $\Box$, though, is not as easy. Adding the KD45 axioms for $\Box$ will result in a sound system, but it is not complete: in particular, the requirement that the intersection between each $\sim_i$-cell and *Act* are singletons is not captured by any of the axioms which have been mentioned here. At this point A proper axiom is not known to the author, and the venture for a complete axiom system for contextual semantic competence will not be investigated further.

7.3 Degrees of Contextual Competence

In chapter 5, formal counterparts to the different types of semantic competence from (Marconi, 1997) were identified. Each such type was characterized relative to one context, i.e. one actual world with epistemic alternatives. In the extended framework, these competence types can be bend in degrees of strength measured by the amount of contexts in which the agent is competent.

As $\mathcal{C}_{\mathsf{2QEL}}^{\Box}$ models include multiple actual worlds, and knowledge in one does not imply knowledge in others by default, agents may in such models be competent with respect to some type in one situation, while simultaneously lacking to be so in another. This allows for a measurement of semantic competence by the quantity

of contexts in which the agent possess one or more kinds of competence. This will shortly be discussed in order to introduce terminology which will allow for a classification of different premises for Frege's Puzzle about Identity to be discussed below. A general hierarchy as in section 5.6 will not be discussed.

S_k Inferential Competence

Modifying the definitions of various competence types from section 5.4 to contextual variants are straight-forward. This can be done merely by relativizing each definition to a context by adding a context parameter in the definitions:

Definition 7.6 (S_k **Knowledge of Co-reference**). Where ω_k is the actual world of context S_k from model M, agent i is said to have S_k *knowledge of co-reference of n and n'* iff

$$M, \omega_k \models_v K_i(\mu(n) = \mu(n')) \tag{7.3.1}$$

◀

Definition 7.7 (**Full S_k Inferential Competence**). Agent i is *fully S_k inferentially competent with respect to n* in model M iff

$$M, \omega_k \models_v (\mu(n) = \mu(n')) \to K_i(\mu(n) = \mu(n')) \tag{7.3.2}$$

for all $n' \in LEX$. ◀.

These two notions of S_k inferential competence can easily be extended to cover competence in multiple contexts: in case an agent is inferentially competent in contexts S_k and S_j, it is said that the agent is $S_k + S_j$ inferentially competent, etc. Where the agent is inferentially competent in all contexts, i.e. is $S_1 + \cdots + S_n$ inferentially competent, the agent is said to *universally have knowledge of co-reference/be universally inferentially competent with respect to n*. This can naturally be expressed by prefixing (7.3.1) with the $\Box$ operators, hence quantifying over all actual worlds:

$$\Box K_i(\mu(n) = \mu(n'))$$

for either n and all co-referring n' or for only n and n'. By (7.2.5), this implies

$$\Box_i(\mu(n) = \mu(n')) \tag{7.3.3}$$

As is the case in regular QEL, formulas with the knowledge operator K_i does not in general allow for existential generalization. As the $\Box_i$ operator can be defined in terms of K_i and $\Box$, it is natural that the lack of generalization also applies in this case. In particular, it is worth noting that (7.3.3) *does not* imply

$$\exists x \Box_i(\mu(n) = x)$$

the importance of which will become evident in the following section.

S_k **Application**

Making Definition 5.15 context-dependent again requires only a small change:

Definition 7.8 (S_k **Application**)**.** Agent i can *apply* name n in context S_k in model M iff

$$M, \omega_k \models_v \exists x K_i(\mu(n) = x) \tag{7.3.4}$$

◀

As for contextual inferential competence, the notion of $S_k + S_j$ application denotes the conjunction of S_k and S_j application, and the notion of $S_1 + \cdots + S_n$ is called *universal application.* Universal application can be captured by

$$\Box \exists x K_i(\mu(n) = x), \tag{7.3.5}$$

as the truth of (7.3.5) requires i to be able to apply n in every context, by the $\Box$ operator.

This notion of competence is the strongest one definable in terms of application in $\mathcal{C}_{2\mathsf{QEL}}^{\Box}$ models. The competence type may be defined in four ways, using each of the two necessity operators and both *de dicto* and *de re* type formulas:

$$\exists x \Box K_i(\mu(n) = x) \tag{7.3.6}$$

$$\Box_i \exists x K_i(\mu(n) = x) \tag{7.3.7}$$

$$\exists x \Box_i K_i(\mu(n) = x) \tag{7.3.8}$$

All four variations are equivalent. From the remarks made in the previous section, it can be seen that both (7.3.6) and (7.3.8) imply, and are implied by

$$\exists x \Box_i(\mu(n) = x) \tag{7.3.9}$$

As the *de re* formulas imply their *de dicto* counterparts, it is further the case that (7.3.6) implies (7.3.5) and (7.3.8) implies (7.3.7). As $\Box_i$ has a wider scope than $\Box$, (7.3.8) implies (7.3.5).

As was remarked in the previous section, (7.3.3) does not imply (7.3.9). This is an important feature of the context models, as it also reflects contextual dissociation of the two competence types. If the implication held, the system would be inconsistent with the empirical findings reported in (Marconi, 1997).

Implicational Aspects As in the mono-contextual case, implicational relationships hold between various conjunctions of competence types. The relationships discussed in section 5.5.2 carry over to the contextual case, when implications within a single context is regarded. However, where the agent possess these different competence types in *different* contexts, this is no longer the case.

In particular, it may be noted that where agent i can apply n in S_k and can apply n' in S_j, this does not imply that i is inferentially competent with n and n' in either context. A counter-example to this implication is presented in the analysis of the following section.

7.4 Frege's Puzzle with Contexts

Everything is now ready to handle the objection from section 6.1.1 to the analysis of Frege's Puzzle in terms of application. The intuitive objection rested on the idea that the astronomers where able to apply names to Hesperus and Phosphorus, but would be informed of their co-reference an identity statement.

In order to set the stage, let M be a $\mathcal{C}_{2QEL}^{\square}$ model with $S = \{S_1, S_2\}$ and actual worlds ω_1 and ω_2. One may regard S_1 as the morning context and S_2 as the evening context from the astronomers example. Include names n_p ('Phosphorus') and n_h ('Hesperus'), and the object v (the planet Venus). Let

$$M, \omega_k \models_v (\mu(n_p) = v) \wedge (\mu(n_h) = v)$$

for $k = 1, 2$ in order for the two names to co-refer in the actual worlds. This is in accordance with the rigidity requirement, (7.1.1) above. Next, add constants for Hesperus, h, and Phosphorus, p, with the requirement that

$$M, \omega_k \models_v (h = v) \wedge (p = v)$$

i.e. that Hesperus, Phosphorus and Venus are in fact the same object. That this identity holds in both actual worlds complies with the requirement of incontingent object identity across actual worlds, viz. (7.1.2) above.

The objection further dictates that the agent is able to identify both Hesperus and Phosphorus in the appropriate situations and in those cases apply the appropriate names. Thus assume that

$$M, \omega_1 \models_v \exists x K_i((p = x) \wedge (\mu(n_p) = x)) \tag{7.4.1}$$

in order for agent i to be able to identify Phosphorus in the morning and apply n_p to it, and

$$M, \omega_2 \models_v \exists x K_i((h = x) \wedge (\mu(n_h) = x)) \tag{7.4.2}$$

for the same to be possible for the agent regarding Hesperus in the evening.

The problematic contradiction in the objection follows from the fact that the astronomers did not know that names n_p and n_h co-refer. Yet, if the further assumption

$$M, \omega_k \models_v \neg K_i(\mu(n_p = \mu(n_h)) \tag{7.4.3}$$

is required, a model can still be constructed satisfying all the above assumptions, hence providing an example showing Millianism consistent with the objection. Such a model is illustrated in Figure 7.4.1.

The model can be constructed by requiring that v, h, p, $\mu(n_h)$ and $\mu(n_p)$ are interpreted as the same object across actual worlds, while requiring that these vary over at least one epistemic alternative. Specifically, v, h and $\mu(n_h)$ should vary in S_1 and v, p and $\mu(n_p)$ should vary in S_2.

Where $W = \{\omega_1, \omega_2, w_1, w_2\}$ with $w_1 \in S_1^i$ and $w_2 \in S_2^i$, an interpretation assigning values as presented in the following table will ensure that M will satisfy (7.4.1), (7.4.2) and (7.4.3) above. This is also illustrated in (Figure 7.4.1)
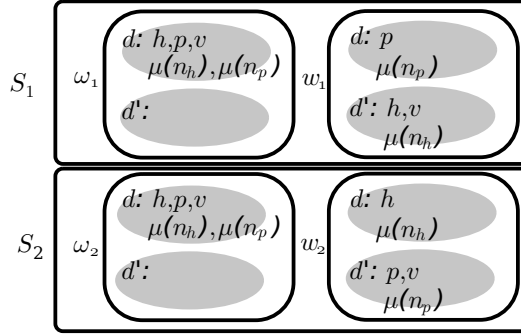
Figure 7.4.1: A $\mathcal{C}_{2QEL}^{\square}$ model where i can apply both n_p and n_h, but is not inferentially competent with regard to the two names.

$\mathcal{I}$	ω_1	ω_2	w_1	w_2
h	d	d	d'	d
p	d	d	d	d'
v	d	d	d'	d'
$\mu(n_h)$	d	d	d'	d
$\mu(n_p)$	d	d	d	d'

In this case,

$$M, \omega_k \not\models_v K_i(\mu(n_p) = \mu(n_h))$$

and hence *an assumption of lacking inferential competence does not result in a contradiction.*

The key point here is that as agent i is not inferentially competent with respect to the two names, i would be informed by being told that the two names co-refer. Further, as the agent is able to identify the referent of each name in the appropriate context, the agent is able to use each name in the appropriate context, or act on information containing it. If, for example, asked to point to Phosphorus in the morning, agent i would be able to do this. Yet, the two names may still differ in cognitive value to the agent, as the agent is not *fully* informed. As the construction shows, this is *not inconsistent with Millianism*.

That the agent can be informed by the true identity statement while being referentially competent in each context stems from the fact that the agent is unable to identify the names as co-referring *across contexts*. For all i knows, it is possible that the two names does not refer to the same object. This is reflected by the model satisfying

$$\Diamond_i(\mu(n_p) \neq \mu(n_h)).$$

From this analysis it can be seen that the objection smuggles into the conclusion more information on the agent's behalf than was introduced in the premises: S_1 application of n_p and S_2 application of n_h was assumed, and it was objected that this lead to a contradiction, since the agent should be informed by the true identity statement $\mu(n_p) = \mu(n_h)$. But for a contradiction to be derived, either S_1 or

S_2 knowledge of co-reference is required. However, such knowledge of co-reference cannot be validly concluded from the assumed premises. Therefore, the problematic contradiction cannot be derived.

The objection has thus been elucidated as a fallacious argument, the fallacy of which consists in using different requirements with respect to what information is available to the agent in premises and what information is available in the conclusion. Hence, the above analysis can be seen as a case of an informational analysis of a problem from the philosophy of language, the structure of which is made clear by focusing on the informational content available to the agent in virtue of his semantic competence.

8 Conclusions and Further Work

As stated in the introduction, this thesis sought to achieve two aims:

1. To construct a formal theory of semantic competence, and

2. To show that the formal theory could be used as an useful analytical tool in uncovering the informational structure behind problems from the philosophy of language.

Both these aims have been achieved.

The first aim was achieved by initially, in chapter 2, outlining a basic theory of meaning for proper names, Millianism, along with the most well-known objections against it, most notably Frege's Dilemma. The simple theory of Millian meaning was taken to be the theory of meaning for which a theory of semantic competence was sought. This endeavor was also commenced in chapter 2: three different philosophical theories of semantic competence were evaluated in order to find one which allowed for an objective, inter-subjective comparison of competence levels. The most viable of the theories considered was the conceptual theory of lexical competence from (Marconi, 1997). Chapters 3 and 4 discussed propositional and quantified epistemic logic from both a philosophical and a modeling perspective, and a general completeness result was shown for many-sorted extensions of QEL. This set the stage for chapter 5, which was dedicated to modeling a simplified version of Marconi's theory. This was done using a two-sorted extension of QEL, where a suitable model-class was defined and a meaning function was added. Based on the results in chapter 4, a sound and complete axiom system was presented, and a logic representing the formal theory had therefore been found. Chapter 5 further presented a validation of the formal theory. It was shown that both the essential ontological properties as well as the competence types from Marconi's theory were present. It was further shown that the formal counterparts of the competence types respected the dissociation of those from Marconi's theory. Thereby, the first aim was accomplished.

To accomplish the second aim, chapter 6 focused on proof on concept. In chapter 6, the model was applied to each of the two disjuncts of Frege's Dilemma. This was done by evaluating the arguments while focusing on the epistemic situation of the agent, i.e. the agents level of semantic competence. It was concluded that once the underlying informational structure of the discussed situations was revealed, neither disjunct proved to be a problem for the Millian theory of meaning. Thereby, the second aim was accomplished.

Yet, an objection was raised to one of the proposed solutions in chapter 6. In order to show that this objection was not fatal for the proposed analysis, chapter 7

was devoted to the construction of a contextual theory of semantic competence. The notion of contexts was incorporated into the models for semantic competence, and the possibilities for finding a complete axiomatic system was discussed, but no completeness result was provided. Therefore, a formal theory, i.e. a logic, for contextual semantic competence was not found. However, the model-theoretic machinery was used to re-analyze the problematic case from chapter 6. It was shown that when the situation was modeled in a contextual model, the epistemic analysis of the disjunct again showed the Puzzle about Identity was in fact unproblematic for the Millian view.

Overall, the constructed formal models for semantic competence have been shown to elucidate informational aspects of the problems posed to the philosophy of language by Frege's Dilemma. In particular, once the informational structure of the problems was clear, it was shown that each argument was far from being as decisive against Millianism as has been the mainstream view in 20th century philosophy of language.

From these conclusions, the present thesis can be seen as supporting two more general lessons with respect to philosophy and mathematics. First, the thesis supports the view that formal logics, and in particular formal non-classical logics such as epistemic logic, can be useful as mathematical modeling tools. As a modeling tool, epistemic logic can be used for *qualitative* modeling of human cognition and cognitive structures, explicitly expressing the informational content of an agent's epistemic situation. This allows for a stringent, transparent analysis of the information content, where the content has tangible structure. As has been illustrated, such models can be used in the analysis of certain philosophical problems, but it is conjectured that such modeling tools will also be useful in the broader field of empirically based research in cognitive science and neuropsychology.

Secondly, the thesis supports the view that formal methods are useful in philosophical analysis. The application of formal methods encourages precision of concepts, and may thus be used to tether intuitions. Providing stringent definitions of used concepts is useful for elucidating nuances which may otherwise be overlooked. In the present case, the first proposed resolution of Frege's Dilemma rested on exactly such insights: the formal notions of semantic competence, meaning and informational content showed exactly why the argument was fallacious. The use of formal models further made it easy to see that the proposed objection included a parameter not accounted for in the models used for the analyses, namely contexts. If guided solely by natural language intuitions, it is not clear whether the cause of the problem nor the solution would have been noticed. This transparency is a direct result of applying formal methods: had Frege had the formal tools available today, the first chapters of many modern textbooks on the philosophy of language may have looked very different from the way they do today.

126

8.1 Venues for Further Research

The formal theory of semantic competence developed through this thesis is a first stab at developing such a theory, but has already been shown useful. Apart from hopefully showing the usefulness of epistemic logical modeling to elements of empirical cognitive science, the formal framework, or extensions of it, is conjectured to be useful in many topics from the philosophy of language and the theory of rational agents and communication. In this final section, some such venues for further research are outlined.

A Logic for Contextual Semantic Competence   An open problem from the present thesis is whether it is possible to find a complete axiomatic system describing the class $\mathcal{C}_{2QEL}^{\square}$. A complete logic for this model class would provide a formal theory for contextual semantic competence.

Extending the Agent Language   As mentioned in section 5.4.2, the two-sorted logic used as a modeling tool is limited in the features it can express. As meaning is assigned by a first-order function, if the agent language was extended to include lexical items for predicates, these could not be assigned an extension. The language is therefore not expressible enough to model semantic competence with respect to anything but the most basic terms from natural language: proper names.

In order to analyze *The Problem of Substitutivity* mentioned in section 2.1.3, or *Kripke's Puzzle about Belief*, cf. (Kripke, 1979), the agent language should be able to express sentences involving intentional contexts, like that of belief. Analyzing the Problem of Substitutivity could lead to conclusions regarding which competence levels allows for the inference, and which do not. Given that Kripke's puzzle is like the Problem of Identity in structure, it is likely that an analysis analogue to that presented here would shed light on which answers the question posed by Kripke should be given, depending on the competence level of the agent in question.

Non-Denoting Terms and Doxastic Attitudes   The present approach has focused on semantic competence solely in the light of *knowledge.* As an intentional attitude, knowledge is very strict, which implies that the types of competence types modeled are stronger than necessary. A focus on weaker operators would give a more fine-grained view on semantic competence. One natural choice would be to weaken the S5 knowledge operator to a KD45 belief operator, or to use the various operators presented in (Baltag and Smets, 2008).

If an approach using weaker modalities is combined with non-constant domain models, this could be used to model individual concepts of non-existing objects. This could possibly shed light on problems of reference to non-existing objects mentioned in section 2.1.3. In particular, such an approach could shed light on the cognitive value of such empty names, and hence on the apparent understanding of such in communicational contexts. Further, as such understanding of proper names will

have no external meaning, a formal modeling of these aspects could very well be put in relation to the semantic internalism/externalism debate, and possibly provide novel insights in this regard.

Multi-Agent Situations   The formal apparatus presented has been defined for multiple agents, yet no multi-agent situations have been analyzed. Extending the formal framework with *announcement operators*, cf. (van Ditmarsch et al., 2008), will result in systems in which one could investigate the effect varying competence levels have on communication in multi-agent settings. Using such an approach, it would be possible to investigate to which degree agents understand one another when their understanding of the language in which they communicate vary. This would further lead to a more natural interpretation of announcement logics, as such do not incorporate the agents' understanding of the language by which the announcements are mediated.

Adding the perspective of what information an agent understands when an announcement is made would also have effects on the type of *actions* agents can perform after being informed. In order to investigate which competence levels are required for which type of action could be evaluated by adding an agent language to a framework like the Logic of Rational Agents, LORA, cf. (Wooldridge, 2000). This framework already incorporates certain communicational aspects, while ignoring the mediating language.

Becoming Competent   Adding announcement operators would further allow investigations of *competence acquisition*. In this respect, it could be investigated which type of knowledge and updates would allow for an agent to become more competent with lexical items. This has informally been discussed in relation to Frege's Dilemma, where it was seen that, under certain circumstances, the agent became inferentially competent after an announcement, cf. e.g. 107. Investigating which kinds of situations and updates allow for an agent to become *referentially* competent with a name is a different problem, and it is conjectured that certain negative results could be shown – in particular, that there exists *no universal update form* which will result in the agent becoming referentially competent with a given word. If this is so, then an interesting problem regarding the human capability of obtaining such competence presents itself.

Formalizing Theories of Reference   The reference relation underlying the meaning function μ has in the present been 'black boxed'. Yet, the type of reference assumed will have an effect on the possibilities for obtaining referential competence. For example, if an indirect theory of reference is assumed, such that the reference of a proper name is determined by some definite description, the negative result mentioned above might dissolve. On the other hand, if a direct theory of reference, like a *causal theory of reference* (Kripke, 1980) invoking a *division of linguistic labor* (Putnam, 1975), is assumed, such a negative result might seem very natural.

128

This aspect could be investigated by adding a branching time aspect and finding an appropriate way to model *reference fixing* by either dubbing or expert opinion. In order to model agents knowledge of the existence of reference fixing experts, the expressibility of *epistemic term-modal logic*, e.g. (Rendsvig, 2010), would be required. The extension of chapter 7 already allows for one way of modeling experts: an agent being able to identify an object in the largest number of contexts can be said to be an expert with respect to that object, and the extension of the appropriate name should therefore be fixed in accordance with this agent's knowledge.

In such a framework, *reference indeterminacy* could be investigated. This could be put in relation to the *meaning indeterminacy* of Dummett, e.g. (Dummett, 2006). In this case, it could be imagined that, as time passes in the model, proper expert knowledge will stop being available, and the name will stop have a determinate reference. In this case, it could be investigated which kinds of competence agents can have with respect to such a name. It could be imagined that the types of competence applicable will be closely related to those of empty names and doxastic competence types.

It is my hope that I will get the opportunity to work on some of these fascinating topics.

Bibliography

Aydede, M. (2010). The Language of Thought Hypothesis, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, fall 2010 edn.

Baltag, A. and Smets, S. (2008). A Qualitative Theory of Dynamic Interactive Belief Revision, *LOFT7*, Amsterdam University Press, pp. 11–58.

Barton, N. H., Briggs, D. E. G., Eisen, J. A., Goldstein, D. B. and Patel, N. H. (2007). *Evolution*, Cold Spring Harbor Laboratory Press.

Blackburn, P., de Rijke, M. and Venema, Y. (2001). *Modal Logic*, Cambridge University Press.

Boer, S. and Lycan, W. G. (1986). *Knowing Who*, Bradford Books / The MIT Press.

Callaway, H. G. (1988). Semantic Competence and Truth-Conditional Semantics, *Erkenntnis* **28, nr.1**: 3–27.

Carlson, L. (1988). Quantified Hintikka-style Epistemic Logic, *Synthese* **74**: 223–262.

Chellas, B. (1980). *Modal Logic: an introduction*, Cambridge University Press.

Collin, F. and Guldmann, F. (2010). *Meaning, Use and Thruth*, Automatic Press / VIP.

Coltheart, M. (2001). Assumptions and Methods in Cognitive Neuropsychology, *in* B. Rapp (ed.), *The Handbook of Cognitive Neuropsychology*, Psychology Press.

Davidson, D. (1984). *Inquiries into Truth & Interpretation*, Oxford University Press, New York.

Devitt, M. (2006). *Ignorance of Language*, Oxford University Press.

Devitt, M. and Sterelny, K. (1999). *Language and Reality: An Introduction to the Philosophy of Language (2nd Edition)*, Blackwell Publishers.

Dummett, M. (2006). *Thought and Reality*, Oxford University Press.

Ebbinghaus, H.-D., Flum, J. and Thomas, W. (1994). *Mathematical Logic*, 2nd edn, Springer.

Fagin, R., Halpern, J. Y., Moses, Y. and Vardi, M. Y. (1995). *Reasoning about Knowledge*, The MIT Press.

Fitting, M. and Mendelsohn, R. L. (1999). *First-Order Modal Logic*, Springer.

Forst, G. (2006). *Noter til kombinatorik of grafteori*, Københavns Universitet, Matematisk Afdeling.

Frege, G. (1980). On Sense and Reference, *in* P. Geach and M. Black (eds), *Philosophical Writings of Gottlob Frege*, Blackwell.

Garson, J. W. (2006). *Modal Logic for Philosophers*, Cambridge University Press.

Grossi, D., Ştefan Minică, Rodenhäuser, B. and Smets, S. (eds) (2010). *Logic and Interactive RAtionality, Yearbook 2010*, Online, see http://www.illc.uva.nl/lgc/seminar/.

Grossi, D., Kurzen, L. and Velázques-Quesada, F. R. (eds) (2009). *Logic and Interactive RAtionality, Yearbook 2009*, Online, see http://www.illc.uva.nl/lgc/seminar/.

Gundersen, L. (2003). Frege: Logik og Sprogfilosofi, *in* P. Lübcke (ed.), *Engelsk og Amerikansk Filosofi*, Politikens Forlag.

Hendricks, V. F. (2006). *Mainstream and Formal Epistemology*, Cambridge University Press.

Hendricks, V. F. and Pedersen, S. A. (2011). *Moderne Elementær Logik, 2. ed.*, Automatic Press / VIP.

Hintikka, J. (1962). *Knowledge and Belief: An Introduction to the Logic of the Two Notions*, Cornell University Press. Reprinted in 2005, prepared by V.F. Hendricks and J. Symons, King's College Publications.

Hintikka, J. (1969). On the Logic of Perception, *Models for Modalities*, D. Reidel Publishing Company.

Hintikka, J. (1984). Are there non-existing objects? Why not? But where are they?, *Synthese* **60**: 251–258.

Hintikka, J. (1994). Different Constructions in Terms of 'Knows', *in* J. Dancy and E. Sosa (eds), *A Companion to Epistemology*, Wiley-Blackwell.

Hintikka, J. (2007). *Socratic Epistemology: Explorations of Knowledge-Seeking by Questioning*, Cambridge University Press.

Hintikka, J. and Sandu, G. (1995). The fallacies of the new theory of reference, *Synthese* **104**: 245–283.

Bibliography

Hintikka, J. and Symons, J. (2007). Visual Identification in Neuroscience, *Socratic Epistemology (Hintikka, 2007)*, Cambridge University Press.

Hughes, G. and Cresswell, M. (1996). *A New Introduction to Modal Logic*, Routledge.

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

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

Loar, B. (2006). Language, Thought and Meaning, *The Blackwell Guide to The Philosophy of Language*, Blackwell Publishing.

Lycan, W. G. (1994). Semantic Competence, Funny Factors and Truth-Conditions, *Modality and Meaning*, Springer.

Lycan, W. G. (2000). *Philosophy of Language: a contemporary introduction*, Routledge.

Lycan, W. G. (2006). Names, *in* M. Devitt and R. Hanley (eds), *The Blackwell Guide to The Philosophy of Language*, Blackwell Publishing.

Marconi, D. (1997). *Lexical Competence*, The MIT Press, Cambridge, MA.

Meyer, J.-J. C. (2001). Epistemic Logic, *The Blackwell Guide to Philosophical Logic*, Blackwell Publishing.

Putnam, H. (1975). The Meaning of 'Meaning', *in* K. Gunderson (ed.), *Language, Mind and Knowledge*, Vol. 7 of *Minnesota Studies in the Philosophy of Science*, University of Minnesota Press, Minneapolis.

Rast, E. H. (2006). *Reference & Indexicality*, PhD thesis, Section for Philosophy and Science Studies, Roskilde University.

Rendsvig, R. K. (2010). Epistemic Term-Modal Logic, *in* M. Slavkovik (ed.), *Procceding of the 12th Student Session*, Online: http://marija.gforge.uni.lu/proceedings.pdf.

Soames, S. (1989). Semantics and Semantic Competence, *Philosophical Perspectives* **3**: 575–596.

Sokal, R. R. (1977). Classification: purposes, principles, progress, prospects, *in* P. Johnson-Laird and P. Wason (eds), *Thinking: Readings in Cognitive Science*, Cambridge University Press, pp. 185–198.

Sosa, E. (ed.) (1992). *A Companion to Epistemology*, Blackwell Publishing.

Stalnaker, R. (1999). Reference and Necessity, *in* B. Hale and C. Wright (eds), *A Companion to the Philosophy of Language*, Blackwell, pp. 534–554.

Tarski, A. (1972). The Semantic Conception of Truth and the Foundations of Semantics, *in* R. M. Harnish (ed.), *Basic Topics in the Philosophy of Language (1994)*, Prentice Hall.

van Benthem, J., Roy, O. and Pacuit, E. (2011). Towards a Theory of Play: A Logical Perspective on Games and Interaction, *Games* **2**(1): 52–86.

van Ditmarsch, H., van der Hoek, W. and Kooi, B. (2008). *Dynamic Epistemic Logic*, Springer.

Wiggins, D. (1999). Meaning and truth conditions: from Frege's grand design to Davidson's, *in* B. Hale and C. Wright (eds), *A Companion to the Philosophy of Language*, Blackwell.

Wikforss, Å. (2009). Are There Understanding-Assent Links?, *in* B. Armour-Garb, D. Patterson and J. Woodbridge (eds), *The Baltic International Yearbook of Cognition, Logic and Communication, Vol. 5: Meaning, Understanding and Knowledge*, Online, see http://thebalticyearbook.org.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*, Harcourt, Brace & company. Project Gutenberg version, released October 22, 2010 (EBook no. 5740), C.K. Ogden's translation.

Wooldridge, M. (2000). *Reasoning about Rational Agents*, The MIT Press.

Zarba, C. G. (2006). *Lecture Notes on Decision Procedures*, Online: http://react.cs.uni-sb.de/teaching/decision-procedures-verification-06/, chapter Ch.1: Many-Sorted Logic, pp. 1–8.