Quantification under Conceptual Covers

Maria Aloni

Quantification under Conceptual Covers

ILLC Dissertation Series 2001-1

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
phone: +31-20-5256090
fax: +31-20-5255101
e-mail: illc@wins.uva.nl

Quantification under Conceptual Covers

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam,
op gezag van de Rector Magnificus
prof.dr. J.J.M. Franse
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de
Aula der Universiteit
op donderdag 25 januari 2001 te 10.00 uur

door

Maria D. Aloni

geboren te Milaan, Italië

Promotor: prof. dr. J.A.G. Groenendijk
Co-Promotor: dr. P.J.E. Dekker

Faculteit der Geesteswetenschappen
Afdeling Wijsgebeerte
Universiteit van Amsterdam
Nieuwe Doelenstraat 15
1012 CP Amsterdam

Printed by PrintPartners Ipskamp

ISBN: 90–5776–0614

Contents

Acknowledgments

First of all I would like to thank my supervisors Jeroen Groenendijk and Paul Dekker. Without Jeroen this thesis would have never been started, and without Paul it would have never been finished.

I would also like to thank my two teachers in Italy Andrea Bonomi and Gennaro Chierchia. The first taught me the basics of formal semantics and, daring to face Italian bureaucracy, coordinated the Erasmus project that brought me to Amsterdam for the first time ten years ago; the second, as supervisor of my master's thesis, contaminated me with his enthusiasm for linguistic research.

In my master's thesis I investigated the issue of the difficult combination of dynamic quantifiers with non-distributive notions such as epistemic modals and presupposition. Quantification under conceptual covers was thought of at first as a solution to these difficulties. The two further applications to the semantics of questions and the logic of propositional attitudes were thought of at a later stage and my colleagues and friends Jelle Gerbrandy and Robert van Rooy deserve a lot of credit here. The first two chapters of this book and myself benefited enormously from the numerous discussion we had on these issues.

I am further indebted to the following people for comments on earlier versions of various parts of this book: Carlos Areces, Raffaella Bernardi, Shai Berger, Rosella Gennari, Peter Krause, Orin Perkus, Martin Stokhof, Frank Veltman, Henk Zeevat, Marc Zuidema and a number of anonymous reviewers.

During the third year of my PhD project I spent two months as a visiting researcher at Stanford University. I would like to thank David Beaver and Brady Clark for the work we did together during that period.

Finally I am very grateful to my friends, my family and mostly Wander for their support during the the last *heavy* months of my PhD project. This thesis is dedicated to my son Lorenzo who contributed to this work by setting, with his arrival, an irrevocable deadline for its delivery. By the way, all typing mistakes are entirely his responsibility.

Amsterdam, December 2000

Prologue

Someone has killed Spiderman. After a careful investigation you discover that John Smith is the culprit and now you want to arrest him. He is attending a masked ball. You go there, but you do not know what he looks like. Is the sentence 'You know who killed Spiderman' true or false in such a situation? On the one hand, the sentence is true, you know that John Smith did it. On the other hand, the sentence is false. Since you do not know what he looks like, you cannot point him out. As far as you know, this person here might be the culprit, or that person there. The evaluation of this sentence seems to be dependent on the way in which the relevant individuals are specified. These can be identified by a number of methods like naming (John Smith, Bill White, and so on) or ostension (this man here, that person there, and so on). If identification by name is assumed, the sentence is true. If identification by ostension is assumed, the sentence is false.

This example illustrates the central idea I defend and investigate in this book. Different methods of identification are operative in different conversational circumstances and the evaluation of fragments of discourse can vary relative to these methods. Classical semantic theory abstracts from the ways in which individuals are identified and therefore has difficulties in accounting for this dependence. The analysis I propose represents different methods of identification and is able to account for their impact on interpretation.

Questions, propositional attitude reports, and quantified sentences containing epistemic modals are examples of linguistic constructions whose interpretation depends on the ways in which objects are given to us. In this thesis I will study these three constructions using the partition theory for questions; modal predicate logic for propositional attitudes; and an intensional dynamic semantics for epistemic modals, respectively. These three theories make crucial use of the notion of a *possible world*. Possible worlds are evaluation points where expressions of the language receive a denotation. In the present context, worlds receive an information-oriented interpretation. A world is meant as representing an epistemic or doxastic possibility, that is, a possible description of what is the case which is compatible with someone's information or belief. The interpretation of questions, propositional attitudes, and epistemic modals crucially involves a shift from one world of evaluation to another. Notions which behave in such a way are usually called *intensional* notions.

The context sensitive constructions that I will consider are classically represented by logical formulae which contain some variable occurring free in the scope of such an intensional operator. In ordinary logical systems, variables are taken to range over bare individuals, and for this reason these systems do not account for the dependence of such constructions on the way in which these individuals are identified.

The analysis I propose maintains the classical representation of this type of sentences, but accounts for their meaning by proposing a non-standard interpretation of variables in intensional contexts. One part of my proposal consists in letting variables range over functions from worlds to objects, rather than over the objects themselves. These functions are traditionally called intensional objects or *individual concepts*, as they formalize (different) ways of identifying objects. The other part consists in making quantifiers range over sets of concepts which (a) are contextually determined and (b) satisfy the following natural constraint: in each world, each individual is identified by one and only one concept in the relevant set. I will call sets of concepts which satisfy this constraint *conceptual covers*. A conceptual cover represents a method of identification. Different conceptual covers represent different ways of looking at one domain. By adopting quantification under conceptual covers in the three previously mentioned theories, the interpretation of questions, propositional attitudes, and epistemic modals are made dependent on the conceptualizations of the universe of discourse which are pragmatically operative. I will show that such a relativization enable us to solve a number of traditional difficulties, and new ones, which emerge in connection with these notions; at the same time we avoid the specific problems which normally arise when we quantify over concepts rather than objects.

Organization of the thesis

The first three chapters of this thesis can be read independently of each other. The chapters 1, 2, and 3 were born as independent articles written in different periods of my graduate studies. Putting independent papers together naturally leads to redundancy and notational inconsistency. I hope that I managed to eliminate most notational variety, but some redundancy was unavoidable. Chapter 4 is meant as a natural compound of each of the previous ones, and has not much sense without them.

Chapter 1 concerns the interpretation of questions and knowing-wh constructions. It has grown out of some material I presented in Leipzig (Sinn und Bedeutung 1998) and Stanford (LLC 1999). In this chapter I present a refinement of the Groenendijk & Stokhof logic of questions which involves relativizing queries to specific conceptualizations of the universe of discourse. I show that in this way a number of difficulties arising for the interpretation of wh-questions and their answers are avoided. I then extend my analysis to two other linguistic theories of questions, the proposition set theory and the structured meaning approach. Part of chapter 1 will appear under the title 'Questions under Cover' in *Proceedings of LLC 8*, edited by D. Barker-Plummer, D. Beaver, J. van Benthem and P. Scotto de Luzio (CSLI publication, Stanford, CA).

In chapter 2, I discuss the interpretation of propositional attitudes, in particular belief reports. The chapter has grown out of Aloni (1998). In the first part, I discuss the classical puzzles arising from the interplay between propositional attitudes, quantifiers and the concept of identity. I compare different reactions to these puzzles in the framework of Modal Predicate Logic and argue in favor of an analysis in which *de re* belief attributions are relativized to the ways of identifying objects used in the specific circumstances of an utterance. In the second part of the chapter, I give this analysis a precise formalization and present Modal Predicate Logic under Conceptual Covers from a model- and proof-theoretic perspective. I compare it with ordinary Modal Predicate Logic and discuss a number of applications.

Chapter 3 discusses the issue of the combination of dynamic quantifiers with 'holistic notions' such as epistemic modality, presupposition and dynamic support. I compare different styles of dynamic quantification, and I argue that all lead to empirical and theoretical difficulties when they are combined with such holistic notions. I then show that quantification under conceptual covers avoids these difficulties. The chapter has grown out of Aloni (1997a) and Aloni (1997b). Most of it has appeared under the title 'Conceptual Covers in Dynamic Semantics' in *Logic, Language and Computation. Volume 3* edited by Patrick Blackburn and Jerry Seligman (CSLI publication, Stanford, CA).

Chapter 4 investigates formal and pragmatic aspects of conceptual covers. After studying a number of formal properties of the notion of a conceptual cover, I compare my identification under conceptual covers with other views of trans-world identification. Next, the pragmatic selection of conceptual covers is discussed. I suggest that the contextual procedures of cover selection are governed by a number of interpretation and generation constraints, which must be *soft*, i.e. violable, in an Optimality Theoretic fashion. I sketch the outline of a Bi-dimensional OT interpretation whose formulation uses concepts from Game Theory. Game Theory turns out to be a promising framework for describing the interplay between the addressee and the speaker in the search for an optimal interpretation of context-dependent natural language expressions.

Chapter 1

Questions

1.1 Introduction

The present chapter concerns the interpretation of interrogative sentences. In particular, the attention will be focused on who-questions and their answers. To fix ideas I adopt as formal framework the partition theory of questions of Groenendijk and Stokhof (G&S). The choice is motivated by the technical sophistication of the G&S system which enables a perspicuous formulation of the problems I intend to discuss. Most of these problems are not peculiar to the G&S analysis though and they trouble (although sometimes in different forms) other approaches as well. This holds in particular for other partition theories of questions (e.g. Higginbotham and May (1981)), but also for proposition set theories (Hamblin (1973), Karttunen (1977)), and structured meaning approaches (e.g. von Stechow (1990), Krifka (1999)).

The structure of the chapter is as follows. In section 1.2, I briefly introduce the G&S logic of questions. In section 1.3, I discuss a number of difficulties that arise for the interpretation of who-interrogatives. I argue that these difficulties are due to the standard method of individuating objects adopted in the G&S analysis. In section 1.4, I propose a modified analysis in which different methods for the identification of objects are available. Identification methods are formalized by what I call 'conceptual covers'. Conceptual covers represent different ways of conceiving the elements of the domain. Questions are then relativized to contextually given conceptual covers. What counts as an answer to a who-question depends on which conceptualizations of the universe of discourse are used in the specific circumstances of the utterance. In the last section, I extend the analysis to other semantic theories of questions.

1.2 Partition Theory

In the partition theory of questions of G&S, the following three principles, formulated at first by Hamblin in the late 50s (see Hamblin (1958)), are formalized:

A To know the meaning of a question is to know what counts as an answer to that question.

B An answer is a statement.

C The possible answers to a question form an exhaustive set of mutually exclusive possibilities.

The meaning of a question is identified with the set of its possible answers (A), that is a set of propositions (B), which determine a partition of the logical space (C).

The formal framework I adopt is based on G&S (1984, 1997). The language under consideration is a language of first order predicate logic with the addition of a question operator '?'.

1.2.1. DEFINITION. [Language] Let PL be a language of predicate logic. The Query Language based on PL is defined as the smallest set QL such that:

1. If $\phi \in PL$, then $\phi \in QL$;
2. If $\phi \in PL$, $\vec{x}$ is a sequence of n variables $(0 \leq n)$, then $?\vec{x}\phi \in QL$;
3. If ϕ and $\psi \in QL$, then $\phi; \psi \in QL$.

Interrogative sentences are obtained by prefixing a question mark and a sequence of n variables to a sentence of PL. A question mark can only occur as outermost operator. We do not have compound interrogatives or quantification into questions, but by the last clause we can form sequences of questions (and assertions). We can distinguish polar questions $(n = 0)$, single-constituent questions $(n = 1)$, and multi-constituent questions $(n > 1)$, as illustrated by the following examples:

$?\exists xPx$	'Did anybody call?'
$?xPx$	'Who called?'
$?x\ x = t$	'Who is Tintoretto?'
$?xyRxy$	'Who invited whom?'

In the partition theory, interrogatives receive an intensional interpretation, and hence a model for our query language will contain a set of possible worlds.

1.2.2. DEFINITION. [Models] A model M for QL is a pair $M = \langle D, W\rangle$ where

(i) D is a non-empty set of individuals;

(ii) W is a non-empty set of worlds w that assign to each individual constant symbol a in PL an item $w(a) \in D$ and to each n-ary relation symbol R of PL a relation $w(R) \subseteq D^n$.

A model is a pair consisting of a set of individuals (the universe of discourse) and a set of worlds. A world is identified with an interpretation function for the non-logical constants in PL. So, a model can be seen as a set of ordinary first order models sharing one and the same domain. It is normal practice in formal semantics to assume models which are large enough to represent the whole space of logical possibilities. I will call such models standard models. A *standard model* for a language QL is a model $M = \langle D, W\rangle$ such that W (the logical space of the model) contains all possible interpretations w of the non-logical constants in PL on D, except, possibly, of the individual constants (singular terms) which can be interpreted as rigid designators in M. A *rigid designator* is a term which denotes one and the same individual in all possible worlds. I will call a model *rigid* if the individual constants are treated as rigid designators. According to the definition above, a rigid model is standard, if it contains all possible interpretations for the predicate constants in the language.

A classical interpretation is assumed for the indicative part of the language. The denotation of an indicative sentence relative to a world is a truth value: $[\![\phi]\!]_{M,w,g} \in \{1, 0\}$ where w is a world and g is a value-assignment to the individual variables in PL.

Interrogatives are analyzed in terms of their possible answers. The denotation of an interrogative in a given world is the proposition which expresses the complete true answer to the question in that world. In what follows I use $\vec{\alpha}$ to denote sequences $\alpha_1, \ldots, \alpha_n$, where the α_i can be variables or individuals.

1.2.3. DEFINITION. [Interrogatives]

$$[\![?\vec{x}\phi]\!]_{M,w,g} = \{v \in W \mid \forall \vec{d} \in D^n : [\![\phi]\!]_{M,v,g[\vec{x}/\vec{d}]} = [\![\phi]\!]_{M,w,g[\vec{x}/\vec{d}]}\}$$

An interrogative $?\vec{x}\phi$ collects the worlds v in which the set of sequences of individuals satisfying ϕ is the same as in the world of evaluation w. If $\vec{x}$ is empty, $?\vec{x}\phi$ denotes, in w, the set of the worlds v in which ϕ has the same truth value as in w. For example, a polar question $?p$ denotes in w the proposition *that* p, if p is true in w, and the proposition *that not* p otherwise. As for who-questions, suppose d_1 and d_2 are the only two individuals in the extension of P in w, then the proposition *that* d_1 *and* d_2 *are the only* P is the denotation of $?xPx$ in w, that is the set of v such that $v(P) = \{d_1, d_2\}$.

Whereas indicatives express propositions, interrogatives determine *partitions* of the logical space. I will write $[\![\phi]\!]_M$ to denote the meaning of a closed sentence[1] ϕ with respect to M. If ϕ is an indicative, $[\![\phi]\!]_M$ is the set of worlds in which ϕ is true. If ϕ is an interrogative, $[\![\phi]\!]_M$ is the set of all possible denotations of ϕ in M. While the meaning of an indicative corresponds to its truth conditions, the meaning of an interrogative is identified with the set of all its possible complete answers. Since the latter is a set of mutually exclusive propositions whose union exhausts the set of worlds, we say that questions partition the logical space. Partitions can be perspicuously visualized in diagrams.

p
$\neg p$

λw [nobody is P in w]
λw [d_1 is the only P in w]
λw [d_2 is the only P in w]
λw[d_1 & d_2 are the only P in w]
$\vdots$
λw[all $d \in D$ are P in w]

In the first diagram, the polar question $?p$ divides the set of worlds in two alternatives, the alternative in which p is true and the alternative in which p is false.

[1] Closed sentences are formulae in which no variable occurs free.

In the second diagram, the single-constituent question $?xPx$ divides the set of worlds in as many alternatives as there are possible denotations of the predicate P within M. Intuitively, two worlds belong to the same block in the partition determined by a question if their differences are irrelevant to the issue raised by the question.

Answers

A question determines a partition of the set of worlds into a number of alternatives. Each of these alternatives corresponds to a complete answer to the question. A (partial) answer is a disjunction of at least one but not all complete answers.[2]

1.2.4. DEFINITION. [Answers] Let ψ and $?\vec{x}\phi$ be closed sentences in QL.

1. ψ is a (partial) *answer* to $?\vec{x}\phi$ in M, $\psi \rhd_M ?\vec{x}\phi$, iff

$$\exists X \subset [\![?\vec{x}\phi]\!]_M : [\![\psi]\!]_M = \cup\{\alpha \mid \alpha \in X\} \neq \emptyset$$

2. ψ is a *complete* answer to $?\vec{x}\phi$ in M, $\psi \;{}_c{\rhd}_M\; ?\vec{x}\phi$, iff

$$[\![\psi]\!]_M \in [\![?\vec{x}\phi]\!]_M$$

Whereas for (non-vacuous) polar questions the notions of a complete and a partial answer collapse, the distinction is non-trivial in the case of constituent questions, for which we can have partial answers which are not complete. For example, both (a) and (b) partially answer (1), but only (b) answers (1) completely.

(1) Who called?

a. Bill did not call.

b. Only Eduard called.

The notion of a partial answer defines the response space generated by a particular query and can be used to characterize the notion of relevance in discourse.[3] Partial answers are replies which exclusively address the issue raised by a question. Complete answers resolve an issue exhaustively. The notion of a complete

[2]These definitions are proposed in G&S (1984) where a more liberal notion of answerhood is also defined, which covers propositions which *imply* rather than *are* complete or partial answers. Here, such over-informative replies do not count as answers.

[3]See Roberts (1996b) and Groenendijk (1999) for examples of such an enterprise. They assume notions of a partial answer though, which are slightly different from the one presented here. Roberts defines partial answers as replies which are incompatible with at least one block in the partition determined by the question (see previous footnote). Groenendijk defines answers in terms of his notion of *licensing*.

answer is usually employed for the analysis of a certain class of question embedding verbs, like *know* or *tell*.[4] These verbs can be analyzed as relations between agents and true complete answers to questions.[5] Roughly, a sentence like 'a knows wh-ϕ' is true iff a stands in the *know*-relation to the denotation of the embedded question, i.e. iff a believes the true complete answer to the question.[6] The notion of a partial answer is not useful for the analysis of embedded interrogatives, as is shown by the contrast between (2) and (3).

(2) If Al knows that Bill did not call, then Al knows who called.

(3) If Al knows that only Eduard called, then Al knows who called.

We expect only the latter to be valid in our logic.

1.3 Methods of Identification

1.3.1 Dilemma

In most analyses of interrogatives a strong link exists between constituent questions and the notion of a rigid designator. On the one hand, a constituent question like 'Who P?' asks for a specification of a set of individuals. This specification requires that these individuals are semantically identified; this means that the terms from which an answer is built up must be rigid designators. On the other hand, an identity question like 'Who is t?' asks for the identification of the denotation of t; this means that if t is rigid, asking 'Who is t?' is a vacuous move. In the G&S logic of questions, two facts make this connection entirely perspicuous. I only state them here postponing a detailed proof to the Appendix A.1.

The first fact says that in a standard model M, the sentence 't is (the only) P' (completely) answers the question 'Who is P?' iff t is a rigid designator in M. (I write $!Pt$ for the sentence expressing the proposition that t is the only P, i.e. $!Pt = \forall y(Py \leftrightarrow y = t)$).

---

[4]These verbs are sometimes called extensional, in contrast to intensional question-embedding verbs (like for instance *wonder*). The former take extensions of interrogatives (that is propositions) as arguments, the latter intensions of interrogatives (that is propositional concepts).

[5]Berman (1991) argues against such an exhaustive analysis of extensional question embedding verbs. See also Ginzburg (1995).

[6]One of the advantages of the G&S analysis is that we can define the notion of a complete answer directly in terms of the denotation of the question, and therefore we have a ready account of embedded uses of questions. Proponents of other approaches have to do some extra work here. See Lahiri (1991), pp. 16-22 for an attempt of a definition of the notion of a complete answer assuming a Hamblin-Karttunen denotation for questions. Her definitions are rather complicated though, and not completely general, as she admits. See Heim (1994) and Krifka (1999) for an analysis of knowing-wh constructions in the Hamblin-Karttunen tradition and in the Structured Meaning framework respectively. Their strategy consists in attempting to match the G&S predictions by complicating the lexical semantics of the relevant embedding verbs (see section 1.5).

1.3.1. Fact. [Rigidity and Answerhood] Let M be a standard model.[7] Then

$$Pt \rhd_M ?xPx \quad \Leftrightarrow \quad t \text{ is rigid in } M$$

$$!Pt \ {}_c\!\rhd_M ?xPx \quad \Leftrightarrow \quad t \text{ is rigid in } M$$

The second fact says that if t is a rigid term, then the question 'Who is t?' is trivial. An interrogative $?\vec{x}\phi$ is *trivial* in M iff a tautology is a complete answer to $?\vec{x}\phi$ in M, i.e., if the partition determined by the question consists of a single block comprising the whole logical space.

1.3.2. Fact. [Rigidity and Triviality]

$$t \text{ is rigid in } M \quad \Leftrightarrow \quad ?x\ t = x \text{ is trivial in } M$$

These two facts have consequences that clash with our intuitions in a dramatic way. Consider again the intuitively valid principle (3) and the standard question-answer pair in (A):

(3) If Al knows that only Eduard called, then Al knows who called.

(A) Who called? ($?xPx$)
Eduard called. (Pe)

Note that according to the facts above, Pe is predicted to partially answer $?xPx$ and (3) is predicted to be valid, only if e is a rigid designator. So we would like 'Eduard' to be rigid. However, if we analyze proper names as rigid designators, then intuitively acceptable identity questions like (4) are rendered vacuous:

(4) Who is Eduard? ($?x\ x = e$)

We are faced with a dilemma: either we have to give up accounting for the 'non-triviality' of (4) or for the correctness of (A) and the validity of (3).

Semantic theories of questions (e.g. Hamblin (1973), Karttunen (1977), Groenendijk and Stokhof (1984)) neglect the first horn of the dilemma. These theories adopt Kripke's influential analysis, according to which proper names are rigid designators.[8] Interrogative sentences are interpreted with respect to standard models

[7] Recall that a standard model is a model containing all possible interpretations under which singular terms are possibly interpreted as rigid designators.

[8] See Kripke (1972), p. 48, Blackwell edition, 1980.

> '[...] although someone other than the U.S. President in 1970 might have been the U.S. President in 1970 (e.g., Humphrey might have), no one other than Nixon might have been Nixon.'

in which names are formalized as constant functions, and as a consequence of this, the acceptability of questions like (4) is not accounted for.

A more information-oriented theory of questions[9] might choose for the second horn. Such a theory would take the strong epistemic connotation of notions like (vacuous) questions and proper answers seriously. Correct (non-vacuous) questions signal gaps in the information of the questioner, and whether a proposition provides an adequate answer depends not only on the content of the proposition, but also on the information state of the questioner. More in particular, questions like (4) are correct because subjects can lack information about the actual denotation of proper names, even though the latter are semantically rigid designators.[10] In order to formalize these intuitions, questions and answers can be interpreted with respect to information states S, which are characterized as subsets of the logical space W in some standard model in which proper names can denote different individuals in different worlds. The non-triviality of questions like (4) can then be captured. In some states (4) is vacuous, in others it partitions the relevant set of worlds in a non-trivial way. On the other hand, it then depends on the relevant information state whether (3) is true, and whether (A) counts as a question-answer pair. This is the case only with respect to states in which e is identified. It should be clear that pursuing this line really means neglecting the second horn of the dilemma, and so giving up the standard account of the correctness of (A) and the validity of (3). The described information-oriented theory fails indeed to define a natural notion capable of discerning standard examples like (A) from strongly marked question-answer pairs like the following:

(B) Is it raining?
I am going to the cinema.

First of all, note that the described information-dependent notion of an answer is not suitable for characterizing standard answers at all. As it stands, it does not even allow you to distinguish the following exemplary pair from (B) above:

(C) Is it raining?
Yes, it is raining.

[9] G&S (1984) already recognized the connection between questions and information (cf. their notion of a pragmatic answer). Jäger (1995), Hulstijn (1997), Groenendijk (1998), Groenendijk (1999) are more recent examples of information-oriented theories of questions. As far as I know though, nobody has explicitly proposed the strategy I describe in what follows.

[10] It is important to notice that the phenomena which are typically considered in discussions of rigid designators (alethic modalities and counterfactuals) are of a different nature than the epistemic phenomena considered by information-oriented theories. Many authors (e.g. Hintikka (1975), Bonomi (1983), Groenendijk et al. (1996)) have distinguished semantically rigid designators from epistemically rigid designators – the former refer to specific individuals in counterfactual situations, the latter identify objects across possibilities in information states –, and concluded that proper names are rigid only in the first sense.

By properly selecting a class of worlds S, any proposition can count as an answer to any question with respect to S. For example, 'I am going to the cinema' counts as an answer to 'Is it raining?' in any S in which it is presupposed that I go to the cinema if and only if it rains. If we want to distinguish standard from marked answers ((C) from (B)), we have to abstract from the particular factual information that can be presupposed in a specific situation, and we must look at the general case. A natural way of doing this involves universal quantification over all possible states. A standard answer is a reply which is an answer with respect to any information state in which the sentence is informative. Note, however, that if proper names may have a non-rigid interpretation, the result of such universal quantification is that no answer to a who-question involving proper names will count as standard. So, no line is drawn between the intuitively correct pair (A) and the strongly marked pair (B): while pair (C) counts as standard, only an information-dependent notion of answerhood holds for both (A) and (B).

To summarize, a dilemma arises for the interpretation of constituent questions and their answers. Either we are unable to account for the correctness of questions asking for the denotation of a term t, or we do not manage to distinguish answers built up from terms t from highly marked situation-dependent answers. Standard semantic theories of questions fail to account for identity questions involving proper names. An information-oriented theory fails to account for standard answers involving proper names. Although it is correct in taking the connection between questions and information seriously and in recognizing the context dependence of the notion of an answer, its treatment of who-question and their answers is inadequate. These constructions are certainly context sensitive, but as the examples in the following section will show, their sensitivity is of a different nature than is captured by an information-dependent notion of an answer.

1.3.2 Context Sensitivity

What counts as a good answer to a question in a given context depends on various pragmatic factors.[11] In this section, I discuss two examples illustrating one specific aspect of this context sensitivity.

Priscilla: Consider the following situation. Your daughter Priscilla is doing her homework. She asks you:

(5) Who is the president of Mali?

In order to give her an adequate answer, you fly to Mali, kidnap Konare (the actual president of Mali), bring him in your living room, and finally utter:

[11] Many researchers have recognized the context-sensitivity of questions and answers: see Boër and Lycan (1985), Ginzburg (1995) and Gerbrandy (2000). In the latter an approach is presented which is close in spirit to mine.

a. He [pointing at him] is the president of Mali.

Unfortunately, by uttering a proposition with, to quote Kaplan, Konare himself 'trapped in it',[12] you have not answered Priscilla's intended question. You better should have said:[13]

b. Konare is the president of Mali.

If there is such a thing as a rigid designator in natural language, the demonstrative pronoun in (a) is it. Still, in the described situation, (a) is not an appropriate answer. Literally providing the actual denotation of the relevant predicate, displaying it concretely, does not always give the required result. The notion of a rigid designator, as it is normally intended, does not seem to cut any ice in relation to these phenomena.

Compare the situation above (call it α) with the following scenario β. You are at a party with many African leaders. Priscilla wants to meet the president of Mali.

(5) Who is the president of Mali?

a. He [pointing at him] is the president of Mali.

b. Konare is the president of Mali.

Assume again that Priscilla does not know what Konare looks like. In context β, (a) is an adequate answer to Priscilla's intended question, and (b) is not. What counts as an answer to a who-question depends on the circumstances of the utterance. In one situation, an appropriate answer consists in giving the name of the man; in another, it consists in pointing out the man himself. Both (a) and (b) can be thought of as providing a characterization of the actual denotation of the relevant predicate. The individual that satisfies the property 'being the president of Mali', namely Konare, is identified in both, only in two different ways: in (a) he is identified by ostension, in (b) by the use of a proper name. Which of (a) and (b) counts as an appropriate answer to (5) depends on which of the two methods of identification is salient in the specific circumstances of the utterance. Since Priscilla, given her purposes, is interested in locating Konare in her perceptual field in context β, an appropriate answer consists in pointing out the man himself. In contrast, given Priscilla's goals in α, identification by proper names is the intended method of identification there. What counts as an answer to who-questions seems to depend on a contextually assumed method of identification.

The following example, which involves knowing-who constructions, illustrates the same point.

[12] See Kaplan (1978), p. 226.

[13] The contrast between (a) and (b) corresponds to the distinction between *real* and *nominal* answers (cf. Belnap and Steel (1976)) or *ostensive* and *descriptive* answer (cf. Hintikka (1976), chapter 3). The analysis I propose here will cover more than just this two-fold distinction.

Spiderman Someone killed Spiderman.

γ: You are at the police department. You have just discovered that John Smith is the culprit. You say:

c. (John Smith did it. So) I know who killed Spiderman.

δ: You now want to arrest John Smith. He is attending a (masked) ball. You go there, but you don't know what he looks like. You say:

d. (This person might be the culprit. That person might be the culprit. So) I don't know who killed Spiderman.

In both contexts, your belief state supports the following information:

(6) John Smith killed Spiderman.

However, only in context γ, (6) intuitively resolves the question:

(7) Who killed Spiderman?

So only in context γ, you can truly utter (c). Again, we find that the G&S logic has difficulties in accounting for these examples. If it is combined with the theory of rigid reference, according to which proper names and demonstratives are rigid designators, it trivially fails. If it is combined with an information-oriented notion of non-rigid reference, it is also inadequate. Intuitively in both examples, the shift from one context to the other does not involve any gain or loss of *information*. In both contexts α and β, Priscilla lacks information about the denotation of 'Konare' and in both γ and δ you don't know what John Smith looks like. Again the difficulty described in this section is not peculiar to the G&S analysis, their system only enables us to give it a perspicuous formulation.

1.3.3 The Flexible Model Strategy

In order to get a handle on the issue, I call *identifiers* in a particular situation the terms that 'belong' to the specific method of identification assumed by the questioner in that situation. For example, in β and δ above demonstratives are identifiers; in α and γ proper names are identifiers. As is evident from the examples in the previous section, natural language terms are identifiers only relative to a particular situation.

One conservative way of modeling this variability consists in formalizing identifiers in the same way as rigid designators, that is, as terms that denote one and the same individual in all elements of a certain model. Their context sensitivity is accounted for by selecting different *models* in different contexts. Paraphrasing the notion of a Flexible Universe strategy discussed in Westerståhl (1984), I call this strategy the Flexible Model (FM) strategy.

Westerståhl (1984) discusses the context sensitivity of determiners in natural language. If talking about your party last night, I say the following sentence:

(8) Everybody was crazy.

I don't mean to attribute madness to everybody on earth, but I clearly refer only to the people at the party. Westerståhl calls these contextually selected domains of quantification *context sets* and argues against the Flexible Universe strategy, which identifies them with temporarily chosen model universes. Although the context dependency I am discussing is somehow different from the one considered by Westerståhl, I will follow the line of his argumentation and I will argue against the Flexible Model strategy which formalizes identifiers as terms denoting constant functions in temporarily chosen (standard) models. Let us consider first how the FM strategy could account for the Priscilla case. In context β (at the party), we can interpret our sentences with respect to a (standard) model M_1 in which demonstratives denote one and the same individual everywhere, while names are analyzed as non-rigid designators. In context α (in your living room), we can adopt a (standard) model M_2 in which proper names denote one and the same individual everywhere, while demonstrative need not. We obtain that (a) counts as an answer to (5) with respect to M_1, and (b) with respect to M_2.

(5) Who is the president of Mali?

a. He [pointing at him] is the president of Mali.

b. Konare is the president of Mali.

Furthermore, the identity questions 'Who is Konare?' and 'Who is he?' can count as non-trivial in M_1 and M_2, respectively. The dilemma seems to disappear: it depends on the model whether the question $?x\ x = t$ is vacuous, or whether the sentence Pt answers the question $?xPx$.[14] By interpreting different types of terms as rigid in different situations, the FM strategy accounts for the variability shown by the examples above within the standard analysis. The right class of identifiers is clearly selected by mechanisms that belong to pragmatics rather than semantics. The FM strategy formally characterizes this selection as the selection of a suitable model. Semantic theory, it is assumed, should simply abstract from these mechanisms.[15] Semantics deals with interpretation conditions, rather than actual

[14]Since we are dealing with temporarily selected *standard* models the difficulty arising for the information-oriented theory is avoided here. Question-answer pairs like (A) or (5) which depend on the assumed method of identification are clearly distinguished from information-dependent pairs like (B) above.

[15]The following quote from Higginbotham (1991), taken from Lahiri (1991), is illustrative of this strategy:

interpretations. It tells you, given a certain model, what is the interpretation of a sentence in that model. Pragmatics determines which model should be assumed in a particular situation in order to obtain the intended interpretation in that situation. By assuming such division of labor between semantics and pragmatics, the FM strategy seems to be able to account for the Priscilla example. Nevertheless, in what follows I will show that such strategy has serious methodological and empirical limitations. Different sets of identifiers should be distinguished also in the semantics if we wish to properly account for the linguistic facts. In order to see this, consider the following situation.

the workshop 1 You are attending a workshop. In front of you you have the list of names of all participants, around you are sitting the participants in flesh and blood. Consider the following dialogue:[16]

(9) A: Who is that man?

B: That man is Ken Parker.

A: Who is Nathan Never?

B: Nathan Never is the one over there.

In this dialogue, we seem to find a shift of identification method. In order to account for it an advocate of the FM strategy would have to adopt two different models depending on which question-answer pair she is willing to interpret. The first pair must be analyzed with respect to a structure in which proper names (and not demonstratives) are interpreted as identifiers. For the second pair we need a model in which demonstratives (and not names) are treated as identifiers. This seems to be methodologically suspect and leads to serious difficulties once we assume a perspective which takes discourses as objects of investigation rather than isolated sentences. Intuitively, (9) is a coherent piece of discourse because no move is a trivial move and each move is consistent with the rest.[17] However, if we assume the FM strategy, the two questions in their non-trivial interpretation do not have any model in common. So we lack a semantic characterization of their compatibility.

'The semantics of questions, as I have presented above, abstracts completely from the ways in which objects, or the things in ranges of higher types that are values of the interrogative variables, are given to us, or may be given to questioner or respondent. It furthermore abstracts from the questioner's motives, if any, and other pragmatic matters. It seems to me that semantic theory, and especially the theory of truth for sentences with interrogative complements, requires this abstraction.'

[16] See Hintikka (1976), p. 56, where a similar example is discussed.

[17] See Groenendijk (1999) for an elegant formalization of such a notion of discourse coherence.

The FM strategy does not only fail on the discourse level, though. The pluralism of identification methods that it allows, is not enough to account for all cases, even on the sentential level, as shown by the next example.

the workshop 2 In the situation above you can ask (10) or assert (11):

(10) Who is who? ($?xy\ x = y$)

(11) I don't know who is who.

A typical answer to (10) is one which specifies a mapping from the set of names to the set of people in the room. In the G&S logic, even if combined with the FM strategy, (10) is trivial and so (11) is contradictory. In order to account for these sentences, we have to improve upon the G&S analysis in which different identification methods cannot play a role simultaneously.

It is interesting to notice that examples involving indexical expressions or demonstratives show the same variability that we find in the workshop cases:

(12) You come with me; you stay here.

(13) Pleased to meet you; pleased to meet you.

There are interpretations of (12) and (13) in which the speaker is not contradicting or repeating herself, because the pronoun 'you' can clearly refer to two different people, even inside a single sentence. After the influential work of Kaplan, the standard way of accounting for indexical expressions involves the introduction of the context as an explicit parameter of the interpretation function. In order to account for cases like (12) or (13), we further have to assume that the contextual parameter, which represents circumstances in continuous change, can assign different values to different occurrences of indexical or demonstrative expressions. In the following section, I adopt the same strategy to account for the variability of the interpretation of who-questions. Different sets of identifiers will be allowed to be selected in different contexts as domain of quantification for different occurrences of wh-expressions.[18] The role of pragmatics in these cases is that of choosing not suitable models, as assumed by the FM strategy, but proper domains of quantification.

[18] See Westerståhl (1984), who proposes to account for the context-sensitivity of determiners by allowing different subsets of the universe to be selected in different contexts as domain of quantification for different occurrences of quantifiers.

Digression A number of researchers (notably Boër and Lycan) have assumed that identity questions can involve predicative uses of the copula rather than equative ones. A question like (10) would then be represented as (a) $?xP\ be(x)(P)$ rather than as (b) $?xy\ x = y$, and would not be trivial under such a representation. Note, however, that such a move would not improve the situation for the simple partition theory. The interpretation of an interrogative like (a) would involve universal quantification over a set of properties, which obviously must be contextually restricted, and, therefore, our semantics would still need to be able to distinguish different sets of properties as possible quantificational domains for different occurrences of the wh-phrase in order to account, for instance, for dialogues like (9). Furthermore, it is not at all clear whether the examples I am discussing here really involve predicational uses of the copula. See Higgins (1973), chapter 5, on this issue. On Higgins' taxonomy of copular sentences, the workshop examples would be classified as identificational rather than predicational, since proper names, demonstratives and *who*, cannot be used predicationally, as shown by the contrast between the first and the last three sentences in the following example:

(14) a. John became fat.

b. John became the president.

c. What did John become?

d. *John became that guy.

e. *John became Bill.

f. *Who did John become?

Higgins (p. 166) further observes: 'Identity sentences [expressed by means of =] are close to identificational sentences, and perhaps if one abstracts from "conditions of use" may be analyzed as identical to them'. The analysis I propose in the following section, maintains the simple representation of identity or identificational questions in terms of logical identity and accounts for their meanings by proposing a non-standard interpretation of identity statements in intensional contexts.

1.4 Questions under Conceptual Covers

In this section, I present a refinement of the G&S semantics in which different ways of identifying objects are represented and made available within one single model. Identification methods are formalized by conceptual covers. Conceptual covers are sets of individual concepts which represent different ways of perceiving one and the same domain. Questions are relativized to conceptual covers. What counts as an answer to a who-question depends on which conceptualizations of the universe of discourse are assumed in the specific circumstances of the utterance.

1.4.1 Conceptual Covers

Conceptual covers are sets of individual concepts satisfying a number of natural constraints. Given a set of worlds W and a set of individuals D, an *individual concept* is any total function from W to D. Concepts represent ways of identifying objects. Examples of concepts are the following:

(a) $\lambda w\ d$ (where $d \in D$);

(b) $\lambda w\ [\text{Konare}]_w$;

(c) $\lambda w\ [\text{the president of Mali}]_w$.

The concept in (a) is a constant function that assigns to all worlds the same value d. The concepts in (b) and (c) assign to each world the individual which is Konare or the president of Mali in that world respectively. I will call the value $c(w)$ that a concept c assigns to a world w, the *instantiation* of c in w.

A conceptual cover is a set of concepts which satisfies the following condition: in each world, each individual constitutes the instantiation of one and only one concept.

Given a set of possible worlds W and a universe of individuals D, a *conceptual cover* CC based on (W, D) is a set of functions $W \to D$ such that:

$$\forall w \in W :\ \forall d \in D :\ \exists! c \in CC : c(w) = d$$

The existential condition says that in a cover, each individual is identified by means of at least one concept in each world. The uniqueness condition says that in no world an individual is counted twice. In a conceptual cover, each individual in the universe of discourse is identified in a determinate way, and different conceptual covers constitute different ways of conceiving of one and the same domain.

Illustration Consider the following situation. In front of you lie two cards. One is the Ace of Spades, the other is the Ace of Hearts. Their faces are turned over. You don't know which is which. In order to formalize this situation, we just need to distinguish two possibilities. The following diagram visualizes such a simple model $\langle D, W\rangle$:

$$\begin{array}{lccc} w_1 & \mapsto & \heartsuit & \spadesuit \\ w_2 & \mapsto & \spadesuit & \heartsuit \end{array}$$

D consists of two individuals ♡ and ♠. W consists of two worlds w_1 and w_2. As illustrated in the diagram, either ♡ is the card on the left (w_1); or ♡ is the card on the right (in w_2).

There are only two possible conceptual covers definable over such a model, namely the set A which identifies the cards by their position on the table and the set B which identifies the cards by their suit:

$$\mathrm{A} = \{\lambda w[\mathit{left}]_w, \lambda w[\mathit{right}]_w\}$$

$$\mathrm{B} = \{\lambda w[\mathit{Spades}]_w, \lambda w[\mathit{Hearts}]_w\}$$

C below is an example of a set of concepts which is not a cover:

$$\mathrm{C} = \{\lambda w[\mathit{left}]_w, \lambda w[\mathit{Hearts}]_w\}$$

Formally, C is not a cover because it violates both the existential condition (no concept identifies ♠ in w_1) and the uniqueness condition (♡ is counted twice in w_1). Intuitively, C is ruled out because it does not provide a proper perspective over the universe of individuals. That C is inadequate is not due to properties of its individual elements, but to their combination. Although the two concepts *the card on the left* and *the Ace of Hearts* can both be salient, they cannot be regarded as standing for the two cards in D. If taken together, the two concepts do not constitute an adequate way of looking at the domain.

1.4.2 Interrogatives Under Cover

I propose to relativize questions to conceptual covers. Contextually supplied conceptualizations determine what counts as possible answers to constituent questions.

I add a special index $n \in N$ to the variables in QL. These indices range over conceptual covers. A model for this richer language QL_{CC} is a triple $\langle D, W, C\rangle$ where W and D are as above and C is a set of conceptual covers based on (W, D). I introduce the notion of a conceptual perspective.

1.4.1. Definition. [Conceptual Perspectives] Let $M = \langle D, W, C\rangle$ be a model for QL_{CC}, and N be the set of indices in QL_{CC}. A *conceptual perspective* $\wp$ in M is a function from N to C.

Conceptual perspectives represent the pragmatic contexts, in that they determine the various identification methods which are used.[19] In order to simplify the notation, I will ignore indices and write $\wp(x)$ for the conceptual cover assigned by $\wp$ to the index of x. Sentences are interpreted with respect to perspectives

[19] The present formalization in terms of conceptual perspectives avoids the issue of how covers are contextually determined. See chapter 4 for a discussion of constraints on the selection of conceptual covers.

$\wp$. Only the interpretation of constituent questions which involves quantification over elements of $\wp$-selected conceptualizations is affected by this relativization. In case of multi-constituent questions, different variables can be assigned different conceptualizations. (Recall that $\vec{\alpha}$ stands for the sequence $\alpha_1, \ldots, \alpha_n$. By $\vec{c}(w)$ I mean the sequence $c_1(w), \ldots, c_n(w)$.)

1.4.2. Definition. [Interrogatives under Cover]

$$[?\vec{x}\phi]^{\wp}_{w,g} = \{v \mid \forall \vec{c} \in \prod_{i \in n}(\wp(x_i)) : [\phi]_{w,g[\vec{x}/\vec{c}(w)]} = [\phi]_{v,g[\vec{x}/\vec{c}(v)]}\}$$

The idea formalized by this definition is that by interpreting an interrogative one quantifies over tuples of elements of possibly distinct conceptual covers rather than directly over (tuples of) individuals in D. If analyzed in this way, a single-constituent question like $?xPx$ groups together the worlds in which the denotation of P is identified by means of the same set of elements of the selected conceptual cover, and a multi-constituent question like $?xyRxy$ groups together those worlds in which the pairs (d_1, d_2) in the denotation of R are identified by means of the same pairs of concepts (c_1, c_2), where c_1 and c_2 can be elements of two different conceptualizations. The following diagram visualizes the partition determined by $?xPx$ under a perspective $\wp$ such that $\wp(x) = \{c_1, c_2, \ldots\}$.

λw [no $c_i(w)$ is P in w]
λw [$c_1(w)$ is the only P in w]
λw [$c_2(w)$ is the only P in w]
λw [$c_1(w)$ & $c_2(w)$ are the only P in w]
$\vdots$
λw[all $c_i(w)$ are P in w]

Due to the definition of conceptual covers, in the first block of this partition no individual in D is P; in the fourth block exactly two individuals in D are P; and in the last block all individuals in D are P.[20]

Illustration Consider a slightly modified version of the card situation described above. In front of you lie two closed cards. One is the Ace of Hearts, the other is the Ace of Spades. You don't know which is which. Furthermore, one of the cards is marked, but you don't know which one. We can model this situation as follows (the dot indicates that the card is marked):

$$\begin{array}{lccc} w_1 & \mapsto & \heartsuit & \spadesuit^\bullet \\ w_2 & \mapsto & \spadesuit & \heartsuit^\bullet \\ w_3 & \mapsto & \heartsuit^\bullet & \spadesuit \\ w_4 & \mapsto & \spadesuit^\bullet & \heartsuit \end{array}$$

Our model now contains four worlds, representing the possibilities which are compatible with the described situation. Now consider two possible conceptual perspectives: $\wp$ and $\wp'$. The former assigns to the index of the variable x the cover that identifies the cards by means of their position on the table, $\wp'(x)$ identifies the cards by their suits:

$$\wp(x) = \{\lambda w[\mathit{left}]_w, \lambda w[\mathit{right}]_w\};$$

$$\wp'(x) = \{\lambda w[\mathit{Spades}]_w, \lambda w[\mathit{Hearts}]_w\}.$$

Consider the following interrogative sentence:

(15) Which card is marked? $(?x\ x^\bullet)$

Example (15) structures the set of worlds in two different ways depending on which perspective is assumed:

$$\text{under } \wp: \quad \begin{array}{|c|} \hline w_1 \\ w_2 \\ \hline w_3 \\ w_4 \\ \hline \end{array} \qquad\qquad \text{under } \wp': \quad \begin{array}{|c|} \hline w_1 \\ w_4 \\ \hline w_2 \\ w_3 \\ \hline \end{array}$$

Under $\wp$, (15) disconnects those worlds in which the marked card occupies a different position. Under $\wp'$, it groups together those possibilities in which the marked card is of the same suit. In other words, in the first case, the relevant distinction is whether the left card or the right card is marked; in the second case the question expressed is whether Spades is marked, or Hearts. Since different

[20] See also fact 1.4.3.

partitions are determined under different perspectives, we can account for the fact that different answers are required in different contexts. For instance, (16) counts as an answer to (15) only under $\wp'$:

(16) The Ace of Spades is marked.

The relevance of this difference is easy to see. Imagine you are playing the following game: you can take a card from the table. If it is the marked card you win one million dollars. In this scenario, given your goals (formalized by perspective $\wp$), (16) does not answer (15).

Consider now the following sentence:

(17) Which card is which card? ($?xy\ x = y$)

Since different variables can range over different covers, we can easily account for examples like (17). Assume $\wp$ assigns different covers to (the indices of) x and y, for instance:

$$\wp(x) = \{\lambda w[\mathit{left}]_w, \lambda w[\mathit{right}]_w\};$$

$$\wp(y) = \{\lambda w[\mathit{Spades}]_w, \lambda w[\mathit{Hearts}]_w\}.$$

If interpreted under such perspective, (17) groups together those worlds that supply the same mapping from one cover to the other, and is not vacuous in our model. The determined partition is depicted in the following diagram:

$$\text{under } \wp : \quad \begin{array}{|c|} \hline w_1 \\ w_3 \\ \hline w_2 \\ w_4 \\ \hline \end{array}$$

The question divides the set of worlds in two blocks: $\{w_1, w_3\}$ and $\{w_2, w_4\}$. The first alternative corresponds to the possible answer (18), the second to the possible answer (19):

(18) The Ace of Hearts is the card on the left and the Ace of Spades is the card on the right.

(19) The Ace of Hearts is the card on the right and the Ace of Spades is the card on the left.

1.4.3 Cardinality

In this section, I show that a natural property of the G&S semantics holds here as well. If two worlds w_1 and w_2 belong to the same block in the partition determined by a who-question, then the same number of (sequences of) individuals satisfy the queried property in w_1 and w_2. The proof of this fact relies essentially on the uniqueness and existence conditions that define conceptual covers. This is a desirable fact which therefore constitutes the main justification for the two conditions.

First of all note that, intuitively, how many-questions and 'numeral answers' seem to be insensitive to methods of identification.[21] Consider again the workshop situation described above, in which two identification methods were equally salient, identification by name and identification by ostension.

(20) How many persons were late today?

This question should determine one and the same partition no matter what perspective is assumed. Example (20) should be analyzed as grouping together those worlds in which the same number of people were late today, regardless of how these are identified.

Secondly, we have strong the intuition that knowing who is P implies knowing how many are P. The following is inconsistent:

(21) I don't know how many people were late today, but I know who was late today.

In our logic these intuitions are satisfied, as can be seen from the following fact:

1.4.3. FACT. [Cardinality] Let M be a model, $\wp$ be a conceptual perspective, g an assignment function and $\alpha \in [\![?\vec{x}\phi]\!]^{\wp}_{M}$ be a block in the partition determined by $?\vec{x}\phi$ in M under $\wp$. Then

$$\forall w, w' \in \alpha : |\lambda\vec{d}\,[\![\phi]\!]^{\wp}_{M,w,g[\vec{x}/\vec{d}]}| = |\lambda\vec{d}\,[\![\phi]\!]^{\wp}_{M,w',g[\vec{x}/\vec{d}]}|$$

[21] It may be useful to notice that although how many-questions are independent of the used *epistemic* identification method formalized by the notion of a conceptual cover, they obviously depend on the presupposed *ontological* method of individuation. Conceptual covers are alternative ways of conceiving one and the same domain and so presuppose an individuation criteria for the individuals in this domain. On the issue of ontological identity and individuation see among others Gupta (1980) and van Leeuwen (1991). Gupta (1980) proposes a formalization of a relativistic view on ontological identity according to which individuation criteria are provided by the meanings of common nouns. Interestingly, he formalizes individuation criteria by means of sets of individual concepts satisfying what he calls the condition of separation, which corresponds to my uniqueness condition on conceptual covers.

Who-questions cannot group two worlds w and w' together, if the sets of (sequences of) individuals which satisfy the queried property in w and w' have different cardinality. If you know the true complete answer to the question 'Who P?', then, fact 1.4.3 says that all worlds in your belief state are worlds in which the predicate P is assigned denotations of equal cardinality. If we assume that how many-questions collect those worlds in which the same number of (sequences of) individuals satisfy the relevant property, we can then conclude that if you believe the true complete answer to the question 'Who P?', then you believe the true complete answer to the question 'How many P?'. In the present logic, you know how many P, if you know who P.

The proof of fact 1.4.3 follows directly from the existential and uniqueness conditions on conceptual covers. Irrespective of which perspective you assume, the number of the sequences of individuals satisfying a certain property doesn't change.[22] If we had allowed questions to quantify over randomly collected concepts, rather than conceptual covers, fact 1.4.3 would have been falsified. As an illustration, consider the model visualized in the following picture:

$$\begin{array}{lll} w_1 & \mapsto & \heartsuit^{\bullet} \quad \spadesuit \\ w_2 & \mapsto & \spadesuit^{\bullet} \quad \heartsuit^{\bullet} \end{array}$$

Consider again the set of concepts C= $\{\lambda w[\mathit{left}]_w, \lambda w[\mathit{Hearts}]_w\}$, which as we saw, is not a conceptual cover. Consider now the following interrogative sentence:

(15) Which card is marked?

Suppose we interpret (15) as grouping together those worlds in our model in which the marked card is the instantiation of the same elements of C. Then (15) would place the two worlds w_1 and w_2 in the same block, thus supplying a counterexample to our cardinality fact. Assume the two worlds constitute your information state. In such a situation, it would be predicted that you know which card is marked without knowing how many cards are marked, which is highly counter-intuitive.

A weaker condition than the one proposed for conceptual covers would be sufficient to prove the cardinality fact 1.4.3. Indeed, it would be enough to relativize questions to sets of concepts satisfying the following requirement: in each world all individuals are the instantiations of the same number of concepts. That is, we could relativize questions to unions of conceptual covers, rather than single ones. I don't have knock-down arguments against this proposal. However, its predictions do not seem to fully match our intuitions. For example, consider again the card situation described in the previous section, depicted by the following

[22] See also the related proposition 4.1.2, in chapter 4.

diagram:

$$\begin{array}{llll} w_1 & \mapsto & \heartsuit & \spadesuit^\bullet \\ w_2 & \mapsto & \spadesuit & \heartsuit^\bullet \\ w_3 & \mapsto & \heartsuit^\bullet & \spadesuit \\ w_4 & \mapsto & \spadesuit^\bullet & \heartsuit \end{array}$$

Consider now the following question:

(22) Which card is marked? $(?x\ x^\bullet)$

Such an analysis would predict that in this situation a question like (22) can have a reading in which it is resolved only after the question 'Which suit is in which position?' has been resolved. Indeed, (22) can be interpreted under the union of the two covers $\{\lambda w[\mathit{left}]_w, \lambda w[\mathit{right}]_w\}$ and $\{\lambda w[\mathit{Spades}]_w, \lambda w[\mathit{Hearts}]_w\}$. On my account, questions like (15) cannot have such a counter-intuitive meaning. If you want to know which card is marked and which suit is in which position, you should express it by asking not one, but two questions. A further counter-intuitive prediction of a theory which allows wh-expressions to range over arbitrary unions of covers is that any constituent question can have readings in which it is resolved only after every other possible issue is resolved. If the domain of quantification is taken to be the union of all conceptual covers (that is the set of all concepts), the partition obtained is indeed extremely fine-grained: each block contains only a single world. On my account, instead, wh-questioners cannot be that demanding. If you want to shift from a state of minimal information to a state of maximal information, you cannot do it by asking a single arbitrary constituent question.

1.4.4 Answers under Cover

We can relativize the notions of a partial and a complete answer to conceptual perspectives in an obvious way. I write $[\![\phi]\!]^{\wp}_M$ to denote the meaning or intension of a closed sentence ϕ in a model M, relative to a conceptual perspective $\wp$.

1.4.4. Definition. [Answers under Cover] Let ψ and $?\vec{x}\phi$ be closed sentences in QL_{CC}.

1. ψ is a (partial) answer to $?\vec{x}\phi$ in M *under* $\wp$, $\psi \rhd_{M,\wp} ?\vec{x}\phi$, iff

$$\exists X \subset [\![?\vec{x}\phi]\!]^{\wp}_M : [\![\psi]\!]^{\wp}_M = \cup\{\alpha \mid \alpha \in X\} \neq \emptyset$$

2. ψ is a complete answer to $?\vec{x}\phi$ in M *under* $\wp$, $\psi\ {}_c\!\rhd_{M,\wp} ?\vec{x}\phi$, iff

$$[\![\psi]\!]^{\wp}_M \in [\![?\vec{x}\phi]\!]^{\wp}_M$$

The dilemma discussed above is solved. On the one hand, problems of identification can be represented as problems of mapping elements from different covers onto each other. It depends on the perspective assumed whether an identity question is a vacuous move.

1.4.5. FACT. [Perspectives and Triviality]

$$?x\ t = x \text{ is trivial in } M \text{ under } \wp \ \Leftrightarrow\ \lambda w\ [t]_{M,w} \in \wp(x)$$

We assume that choices of $\wp$ which render questions vacuous are ruled out by general conversational principles.[23] We can thus account for the fact that 'Eduard is Eduard' hardly counts as an adequate answer to 'Who is Eduard?'. The latter, if genuine, asks to map the concept *Eduard* to an element of a conceptualization which crucially does not include it.

On the other hand, terms from which answers are built up need not be rigid designators. It suffices that their interpretations are elements of the assumed methods of identification. We can then account for the difference between the question-answer pairs in examples (A), (B) and (C) in section 1.3.1:

1.4.6. FACT. [Perspectives and Answerhood] Let M be a standard model and $p \not\equiv q$. Then

(A) $\forall\wp:\ \lambda w\ [t]_{M,w} \in \wp(x)\ \Leftrightarrow\ (!)Pt\ {}_{(c)}\!\rhd_{M,\wp}\ ?xPx$;

(B) $\forall\wp : q\ {}_{(c)}\!\not\rhd_{M,\wp}\ ?p$;

(C) $\forall\wp : p\ \ {}_{(c)}\!\rhd_{M,\wp}\ ?p$.

'Eduard called' counts as an appropriate answer to 'Who called?' if and only if the interpretation of 'Eduard' is part of the assumed conceptual cover (cf. (A)). Although they are context dependent, who-questions and their answers are clearly distinguished from highly marked pairs such as 'Is it raining? - I am going to the cinema'. The adequacy of the former depends on the perspective assumed, the correctness of the latter relies on the factual information presupposed (cf. (B)). Standard pairs such as 'Is it raining? - It is raining' are always correct irrespective of the circumstances of the utterance (cf. (C)).

Finally, notice that our analysis allows the characterization of a notion of knowing-who that is relative to a perspective $\wp$: a sentence like 'a knows $?\vec{x}\phi$' is true in w under $\wp$ iff a stands in the *know*-relation to $[\![?\vec{x}\phi]\!]^{\wp}_{w}$ in w, i.e. iff a believes the true complete answer to the question under $\wp$. In this way, we can account for the context sensitivity of knowing-who constructions illustrated by the Spiderman case discussed in section 1.3.2.

1.5 Other Semantic Theories of Questions

In this section, I show that the present analysis does not only apply to the G&S theory of questions, but can also be exported to other frameworks.[24] In particular, I will consider the proposition set theory (section 1.5.1) and the structured

[23] See chapter 4.

[24] See Ginzburg (1996), Higginbotham (1996), and Groenendijk and Stokhof (1997) for recent and exhaustive overviews.

meaning theory (section 1.5.2). In the last section, I discuss Ginzburg's pragmatic analysis and compare my account with his explanation of the fact that different contexts require different answers.

Recall Hamblin's three postulates presented at the beginning of section 1.2, repeated here:

A To know the meaning of a question is to know what counts as an answer to that question.

B An answer is a statement.

C The possible answers to a question form an exhaustive set of mutually exclusive possibilities.

All of the semantic analyses of questions I will consider satisfy the first principle. A theory of questions should provide an account of the answers that an interrogative allows, and this is obtained by relating the meaning of questions to the meaning of their answers.[25] As we saw, approaches that interpret questions as determining **partitions** of the logical space (e.g. G&S (1983, 1997), Higginbotham and May (1981))[26] define the meaning of questions in terms of their *complete* answers and therefore accept all of the three Hamblin postulates given above. The two theories I am going to discuss, the proposition set theory and the structured meaning approach, assume different notions of a possible answer as primary, and so reject one or the other of the Hamblin principles, viz. C and B respectively.

1.5.1 Proposition Set Theory

The sets of propositions theory goes back to Hamblin (1973)[27] and Karttunen (1977), and constitutes one of the most influential theory among the linguistic account of questions. Such an analysis states the meaning of an interrogative in terms of the meaning of what I will call its *singular positive* answers. Singular positive answers are answers that fill in a referential constituent for the *wh*-word in the question and do nothing else. According to these theories, questions are sets of propositions (postulate B), representing the possible singular positive answers to the question (postulate A). However, the latter do *not* form an exhaustive set of mutually exclusive possibilities (postulate C is rejected). In the analysis of

[25]See Hintikka (1976) for an example of a theory which does not define the meaning of interrogatives in terms of the meaning of their answers.

[26]See Groenendijk and Stokhof (1997), pp. 1103-1104, for a comparison between their analysis and Higginbotham and May (1981).

[27]As noted by Groenendijk and Stokhof (1997), Hamblin (1973) does not satisfies the postulates formulated in Hamblin (1958).

Karttunen,[28] questions are analyzed as follows (recall that by $\vec{\alpha}$ I mean sequences $\alpha_1, \ldots, \alpha_n$):[29]

1.5.1. Definition. [Karttunen]

(i) yes-no question:

$$_K[\![?\phi]\!]_{M,w,g} = \{p \mid (p = [\![\phi]\!]_M \vee p = [\![\neg\phi]\!]_M) \;\&\; w \in p\}$$

(ii) constituent question:

$$_K[\![?\vec{x}\phi]\!]_{M,w,g} = \{\lambda v\; [\![\phi]\!]_{M,v,g[\vec{x}/\vec{d}]} \mid \vec{d} \in D^n \;\&\; [\![\phi]\!]_{M,w,g[\vec{x}/\vec{d}]} = 1\}$$

The denotation of a question in a world is identified with a set of propositions, rather than with a set of worlds as in the G&S analysis. An interrogative like 'Did John call?' denotes in world w the set containing the proposition *that John called*, if John called in w or *that John did not call* otherwise. An interrogative like 'Who called?' denotes in w the set consisting of those propositions which are true in w and which say of some individual that (s)he called: {*that John called, that Mary called, that Bill called,* ...}. While in the G&S logic, meanings of constituent questions were characterized as a set of mutually exclusive propositions, here their denotations are identified with a set of mutually compatible possibilities.[30]

[28] The difference between the Hamblin and the Karttunen analysis is that according to Hamblin, interrogatives denote the set of all their possible singular answers, whereas Karttunen takes them to denote the set of all *true* singular answers. The reason of this modification has to do with the interpretation of (a) sentences like 'Who is elected depends on who is running', which can be paraphrased as 'the *true* answer to the former question depends on the *true* answer to the latter question'; and (b) question embedding verbs like *tell*, *indicate*, which become factive if they take a question complement.

[29] Here I consider only who-questions and ignore which-questions. For these the Karttunen analysis differs from the G&S analysis on a further aspect. See footnote 33.

[30] It has been argued that this feature of the set of proposition theory can be useful to account for the so-called *mention-some* interpretations of constituent questions. In certain situations, interrogatives do have a number of fully adequate and mutually *compatible* answers. Consider the following sentence uttered by a tourist in Amsterdam (this example is due to G&S):

(23) Where can I buy an Italian newspaper?

Intuitively (23) can be *completely* resolved by an answer that mentions *some* places in which Italian newspapers are sold, thus a positive (possibly plural) answer, rather than a complete answer, as predicted by the G&S analysis, which provides a full specification of *all* such places. Note, however, that the contrast between mention-some and mention-all interpretations is still not totally understood and it seems that the Karttunen theory would need non-trivial amendments in order to properly account for it. See van Rooy (1999) and van Rooy (2000b) for a recent and interesting analysis of these phenomena. Robert van Rooy's account is based on the assumption that questions are asked in order to resolve decision problems. This view gives an explanation of the fact that certain types of responses can intuitively resolve a question like (23), although the standard partition semantics approach predicts that they cannot.

Since many of the examples we have considered in this chapter involve embedded uses of interrogative sentences,[31] let us see now how these can be treated in the proposition set theory. Karttunen proposes the following analysis of knowing-wh constructions: in order for a sentence like 'a knows Q' to be true in w, the subject has to believe in w every proposition in the denotation of the embedded question in w. This analysis leads to a number of counter-intuitive results[32] the most striking one is that a sentence like 'John knows whether Bill called' would not be entailed by a sentence like 'John knows who called'. In order to remedy these inadequacies, Heim (1994) proposes the following amendment for the lexical semantics of *know* in a proposition set framework: a sentence like 'a knows Q' is true in w iff a believes in w the proposition $\lambda w'[_K[\![Q]\!]_{M,w'} = {}_K[\![Q]\!]_{M,w}]$. In order for a wh-knowledge attribution to be true in w, the subject's belief state must contain only worlds in which the embedded question receives exactly the same Karttunen denotation as in w. For instance, 'John knows who called' is true iff John's belief state is a subset of the set of worlds in which the same set of individuals called as in the world of evaluation, that is, iff John believes the true complete answer to the question. For who-questions, Heim's proposal manages to match the correct predictions of the partition theory,[33] therefore we will assume it in the following discussion.

It is easy to see that this Karttunen-Heim analysis of questions leads to precisely the same difficulties as the G&S theory in connection with the phenomena discussed in this chapter.

First of all, the dilemma presented in section 1.3.1 arises here as well. Clearly, on Karttunen's account, only replies employing rigid terms can count as answers to constituent questions and identity questions concerning rigid terms are trivial, since their only possible answers are tautologies. Recall that the notion of an answer directly definable here is that of a positive singular answer:

1.5.2. DEFINITION. [K-Answers] ψ is a (singular positive) *answer* to $?\vec{x}\phi$ in M, $\psi \,{}_K\!\rhd_M ?\vec{x}\phi$, iff

$$\exists w, g : [\![\phi]\!]_M \in {}_K[\![?\vec{x}\psi]\!]_{M,w,g}$$

[31] In particular, the choice of taking conceptual covers as domains of quantification for wh-expressions rather than randomly collected sets of concepts are better motivated in connection to embedded uses of questions whose interpretation involves a universal quantification over these concepts (recall the arguments in section 1.4.3).

[32] See Groenendijk and Stokhof (1984) for a detailed discussion of these difficulties.

[33] The match is not perfect though and problems arise in connection with which-questions, as noted by Heim herself, who proposes an enriched variant of the Karttunen analysis as a solution which employs structured propositions. I will disregard this issue, which, although interesting, is not relevant to the main theme of this work. Note that the interpretation of which-questions constitutes a problem for most existing approaches, which fail to account properly for their 'asymmetric nature' (e.g. 'Which men are bachelors?' $\neq$ 'Which bachelors are men?').

The notion of a trivial question can be characterized in Karttunen's theory as follows: a question is trivial in $M = \langle D, W\rangle$ iff its denotation in some w in W is the singleton set containing the trivial proposition W. The following two facts clearly hold:[34]

1.5.3. Fact. [K-Answerhood and Rigidity] Let $M = \langle D, W\rangle$ be a standard model.

$$Pt\ {}_K{\rhd}_M ?xPx \quad \Leftrightarrow \quad t \text{ is rigid in } M$$

1.5.4. Fact. [Rigidity and Triviality] Also in K:

$$t \text{ is rigid in } M \quad \Leftrightarrow \quad ?x\ x = t \text{ is trivial in } M$$

Furthermore, the Karttunen-Heim analysis, as it stands, simply ignores the various pragmatic factors that play a role in determining what counts as a good or a bad answer in different contexts. Therefore, it cannot account for the Konare or Spiderman cases, nor for the variability illustrated by the workshop examples. As a remedy to this, we can adopt the same strategy we adopted for the G&S analysis, namely we make wh-expressions range over elements of conceptual covers rather than over individuals of the domain. As above, special indices ranging over covers are added to the variables in the language, whose value is determined by conceptual perspectives $\wp$. The interpretation of wh-interrogative sentences is relativized to these perspectives:

1.5.5. Definition. [Karttunen under Cover]

$${}_K[\![?\vec{x}\phi]\!]^{\wp}_{M,w,g} = \{\lambda v\ [\![\phi]\!]_{M,v,g[\vec{x}/\vec{c}(v)]} \mid \vec{c} \in \prod_{i\in n}(\wp(x_i))\ \&\ [\![\phi]\!]_{M,w,g[\vec{x}/\vec{c}(v)]} = 1\}$$

From this definition, assuming Heim's semantics for *know*, we obtain a perspective relative notion of knowing-who constructions and thus we can account for the Spiderman case: 'a knows Q' is true in w under $\wp$ iff a believes in w the proposition $\lambda w'[{}_K[\![Q]\!]^{\wp}_{M,w'} = {}_K[\![Q]\!]^{\wp}_{M,w}]$.

The alternative possibility of letting wh-expressions range over arbitrary sets of salient concepts, rather than conceptual covers would lead here to the same counter-intuitive results discussed in section 1.4.3 for the relation between knowing-who and knowing-how-many constructions.

[34] The proof of the first fact is in the Appendix. The proof of the second fact is trivial.

1.5.2 Structured Meaning Theory

The Structured Meaning or Categorial Theory of questions originates in the work of Ajdukiewicz in the late 20s and was further developed by among others Tichy (1978), Hausser and Zaeffer (1979), von Stechow and Zimmermann (1984), von Stechow (1990) and more recently Krifka (1999). Such a theory analyzes questions in terms of their answers (postulate A). It does not assume, however, that answers do belong to a uniform category (postulate B which states that answers are sentences is rejected). Sub-sentential answers, often called *constituent* or term answers, play a crucial role in the structured meaning analysis. The category of an interrogative is chosen in such a way that in combination with the category of its constituent answers it yields the category of indicative sentences. As a result different kinds of interrogatives are of different categories and semantic types. In Krifka (1999), questions are defined as functions that when applied to the meaning of the possible constituent answers, yield the meaning of the corresponding full sentential answers. Polar questions expect 'yes' or 'no' as constituent answers, which are analyzed as propositional operators of type $\langle t, t\rangle$[35] which retain or reverse the truth value of the proposition respectively. Yes-no questions are then expressions of type $\langle\langle t, t\rangle, t\rangle$. For example:

(24) Q: Is it raining? $\lambda f[f(p)]$

A: No. $\lambda q[\neg q]$

Q(A): It is not raining. $\lambda f[f(p)](\lambda q[\neg q]) = \neg p$

Instead, single wh-questions take singular noun phrases as constituent answers, whose semantic type is e, that of expressions referring to entities.[36] They are then assigned the type $\langle e, t\rangle$. For example:

(25) Q: Who called? $\lambda x[P(x)]$

A: Mary. m

Q(A): Mary called. $\lambda x[P(x)](m) = P(m)$

One advantage of the categorial approach over theories which analyze questions in terms of sets of propositions has to do with the interpretation of alternative polarity questions (see Krifka (1999)).[37] Both the Karttunen or the G&S

[35] Krifka assumes an extensional type theory. In an intensional type theory they would be of type $\langle\langle s, t\rangle, t\rangle$.

[36] Sometimes the type of a generalized quantifier is assumed for constituent answers to wh-questions.

[37] It has been argued that a further advantage of a structured meaning approach is that it allows a straightforward account of the information structure in answers (see again Krifka

analysis cannot distinguish between the following two questions which are assigned the same denotation if the former is interpreted as a polarity question. A categorial approach, instead, can account for the fact that they allow different answers:

(26) Q: Do you like Milan, or don't you? $\lambda p[p = L(m, y) \vee p = \neg L(m, y)]$ (type: $\langle t, t\rangle$)

A: I don't. $\neg L(m, y)$ (type: t)

(27) Q: Do you like Milan? $\lambda f[f(L(m, y))]$ (type: $\langle\langle t, t\rangle, t\rangle$)

A: No. $\lambda q[\neg q]$ (type: $\langle t, t\rangle$)

Less straightforward, however, is the structured meaning analysis of embedded questions. The fact that interrogatives are not assigned a uniform semantic type requires the assumption of special lexical rules for the various question embedding verbs, since the latter can take (coordinations of) interrogatives belonging to different categories. Krifka (1999) proposes the following entry for *know*, which reduces knowing-wh to knowing-that constructions: a knows Q iff for every full answer A that answers Q, if A is true then a knows that A, if A is not true then a knows that not A. Recall that full answers are the result of the application of the question to the term answers. For example, 'Mary called' is a full answer to the question 'Who called?'.

It is easy to see that the difficulties emerging for the G&S and the Karttunen analyses arise in the structured meaning theory as well, although in a different form. Note first of all that all expressions belonging to the right category (e.g. all replies in (28b), which are of the required type e) seem to be accepted as answers, so there is no strict link between rigidity of terms and the notion of answerhood:

(28) a. Who is the president of the United States?

b. - Bill Clinton.

- That guy [pointing at Clinton].

- Hillary's husband.

- Donald Duck.

- That guy [pointing at Donald Duck].

- Uncle Scroodge's nephew.

(1999)). Note, however, that the topic-focus structure in answers can be accounted also by analyses which assume a proposition set or a partition interpretation for questions. See for example Roberts (1996b) which assumes a Rooth style representation of focus (cf. Rooth (1992)) and a proposition set analysis of questions; and the recent Aloni et al. (1999) which interprets intonation in terms of presupposition of topics under discussion and questions (or explicit topics) as domain restrictions which uniquely determine partitions of the logical space.

A first consequence of this is that the categorial approach, as it stands, cannot explain why different circumstances require different answers employing different methods of identification for the objects in the universe (see the Priscilla case). A further consequence has to do with embedded uses of questions. Note that on Krifka's account, as far as I understand it, in order for the following sentence to be true:

(29) Priscilla knows who the president of the United States is.

Priscilla would have to know, for every full answer A to (28a), that A is true, if A is true, and that A is false, otherwise. Thus (29) would be judged false in a situation in which Priscilla knows that Bill Clinton is the president, but cannot recall whether he is Hillary's husband or Uncle Scroodge's nephew. The problem is that *know* is taken to quantify, universally, over the set of all structurally acceptable answers, while it should clearly quantify only over *good* answers.[38] The categorial approach, as it stands, clearly lacks a characterization of what counts as a good answer to a question in a specific context. The most straightforward way to amend it consists in letting (multi-)constituent questions take as constituent answers (sequences of) expressions of the category of individual concepts (of type $\langle s, e \rangle$) rather than entities (of type e) and then let the contextually operative conceptual covers restrict the domain of application of the function expressing the question meaning. Again we add special CC-indices to the variables of the language and let conceptual perspectives $\wp$ determine their value. A sequence of singular terms $t_1, \ldots, t_n$ will be a *good* term answer to a (multi-)constituent question $\lambda \vec{x} \phi({}^{\vee}\vec{x})$ under $\wp$ iff for all $i \in n$, the interpretation of t_i in the model is an element of $\wp(x_i)$. Good full answers result from applying questions to good constituent answers. We then obtain that John knows Q under $\wp$ iff for every good full answer A to Q under $\wp$, if A is true then John knows that A, if A is not true then John knows that not A. Again, the choice of taking conceptual covers rather than arbitrary sets of concepts is crucial, in particular for the interpretation of knowing-who constructions which involve universal quantification over these concepts (see again the arguments in section 1.4.3).

1.5.3 Pragmatic Theory

Recently, the issue of the context sensitivity of questions has received new attention in the linguistic literature in the work of Jonathan Ginzburg. In order

[38]Krifka observes that 'non-exhaustive' interpretation of *know* (see Berman (1991)) can be captured in his analysis, by positing that knowing *some* answers to Q may be sufficient in order to know Q in certain circumstances, with the universal quantification over answers being just the default option. Note, however, that the cases I am discussing here do not involve non-exhaustive interpretations in the sense of Berman (1991), but rather exhaustive interpretations which intuitively involve universal rather than existential quantification, but not over all structurally acceptable answers, but over a specific subset of them.

to account for the influence of pragmatic factors on the interpretation of questions and answers, Ginzburg (1995) proposes what he calls a *relative* notion of an answer resolving a question. In his theory, questions are analyzed as in the structured meaning approach (though in the framework of situation semantics, rather than possible world semantics); extensional question-embedding verbs are analyzed as imposing a resolvedness condition on their interrogative complement and what counts as resolving crucially depends on contextual parameters, such as the goals and inferential capabilities of the questioner. The Spiderman case discussed above could then be analyzed roughly as follows. Recall the relevant situation. Someone killed Spiderman. In context γ, you are at the police department investigating the murder. In context δ, you are at a ball with the intention to arrest the culprit. In the two contexts, you are after two different goals. A goal can be described by a proposition, intuitively, the proposition that is true once the goal desired by the relevant agent is fulfilled. Going back to the Spiderman example this means:

goal in γ = You know the name of the culprit.

goal in δ = You know what the culprit looks like.

In the situation described, (30) below is true in both contexts. The proposition expressing the goal in γ can be 'inferred' from (30), but the one expressing the goal in δ cannot. So, (31) is true in one context and false in the other.

(30) You know that John Smith killed Spiderman.

(31) You know who killed Spiderman.

However, once we try to formalize this analysis we get into problems. According to the theory of rigid reference, which Ginzburg seems to adopt, the following two propositions are still equivalent:

(32) John Smith killed Spiderman.

(33) He [pointing at John Smith] killed Spiderman.

Hence it is not clear how (30) and (34) can have different implications:

(34) You know that he [pointing at John Smith] killed Spiderman.

The simple introduction of goals and perspectives as explicit parameters of the answerhood relation is not sufficient to explain the phenomenon I discussed here and needs to be combined with a more sophisticated account of how objects are identified in cognitive states. Identification under conceptual covers gives the required ingredients for such an account.

Furthermore, in this chapter, I have also showed a different way of formalizing the same idea that goals and perspectives are relevant for an analysis of questions. On Ginzburg's approach, different answers resolve, in different contexts, an interrogative whose meaning stays constant. In my analysis, an interrogative expresses different partitions in different contexts, because in different contexts different domains of quantification are selected for the wh-expression. In Ginzburg's theory, goals and perspectives are parameters of the answerhood relation, here they play a role in selecting a domain of quantification. Consider Ginzburg's argument in favor of his division of labor (see Ginzburg (1995), pp. 170-171). He distinguishes two ways in which misunderstanding can arise in a dialogue, one arising from failure to communicate a content, the other from failure to share a 'perspective'. He uses the following examples to illustrate this contrast:

(35) a. A: John left.
B: No, look, he's sitting outside, chatting up Milena.
A: I meant John Schwitters.

b. A: Everyone will support the decision.
B: Including the CS people?
A: I meant every linguist.

(36) A: [yawns] That was a boring talk.
B: No, it wasn't.

In (35) B fails to grasp a content. In (36) the participants fail to share a 'perspective'. An exchange such as (36) can be explained 'by saying "boredom is relative", that is, by positing that the relation denoted by the use of "boring" in (36) has an additional 'perspectival' argument which is filled by, say, a mental state of an agent. In contrast to (35), where the impasse in the conversation can be repaired by agreeing on a single way to fix the contextual parameter (establishing what the domain of quantification is), there is no such requirement with (36) where the disagreement can be patched up with each conversationalist still holding on to his perspective' (Ginzburg (1995), pp. 170-171). Ginzburg then argues that the context dependency of questions and answers resembles more the perspectival mismatch in (36) than the semantic mismatch in (35), and hence is better captured by assuming a relative notion of resolvedness rather than by adopting the domain-selection strategy. My claim, in relation to the phenomena I have been discussing in this chapter, is exactly the opposite. Although some of the examples of context sensitivity discussed by Ginzburg may be cases of perspectival mismatch, the cases of dependence on the method of identification that I have considered here are more similar to cases of semantic mismatch. As an illustration, observe the close resemblance between the dialogues in (35) and the following dialogue:

(37) A: Who is the president of Mali?

B: That [pointing at someone] is the president of Mali.

A: I meant, what is his name.

In this example, as in (35) and in contrast with (36), the misunderstanding between A and B is totally resolved by A's remark and it would be weird for B to hold on her own perspective. Although a certain degree of vagueness and 'perspectival' factors might play a role in connection to the general notion of what counts as a good answer to a question in a specific context, I conclude that the cases I have been discussing in this chapter are better captured by assuming that different contexts select different quantificational domains for the wh-phrase, rather than by simply positing an additional contextual or perspectival argument on the notion of resolvedness. A clear advantage of the former strategy is that it allows a straightforward account of the fact that the described context sensitivity is specific of constituent questions and their answers,[39] an analysis of these cases in the style of Ginzburg's would lack a natural explanation of this fact. Furthermore, the context dependency of quantificational domains is a pervasive phenomenon in natural language, and it is not surprising that it arises with wh-expressions as well. Recall, however, that the domain selection I discuss here is somehow different from the one normally recognized (see Westerståhl (1984)). The context does not only decide on which set of objects is salient at the moment of the utterance, but also on which ways of identifying the salient objects are relevant and should be taken into account.

1.6 Conclusion

A domain of individuals can be observed from different angles. Our interpretation of who-questions and their answers may vary relative to which ways of identifying objects we assume. By letting wh-expressions range over elements of conceptual covers, we can account for this variability while maintaining the intuitive characterization of constituent questions as asking for the specification of a set of determinate individuals. The elements of a cover do not stand for representations of individuals but rather for the individuals themselves but identified in a particular way.

[39] The context dependency of polar questions and their answers is of a different kind and can be better captured by an information-dependent notion of an answer (see section 1.3.1).

Chapter 2

Belief

Consider the following sentences from Quine:

(38) a. Philip is unaware that Tully denounced Catiline.
b. Philip is unaware that Cicero denounced Catiline.
c. Philip is unaware that x denounced Catiline.

Suppose (a) is true and (b) is false. What is the truth value of (c) under the assignment that maps the variable x to the individual which is Cicero and Tully?[1]

After 50 years, Quine's question is still puzzling logicians, linguists and philosophers. In the present chapter, this and other paradoxes of *de re* propositional attitude reports are discussed in the framework of modal predicate logic. In the first part of the chapter, I compare different reactions to these paradoxes in such a framework and I argue in favor of an analysis in which *de re* propositional attitude reports are relativized to the ways of identifying objects used in the specific circumstances of an utterance. The insight that different methods of identification are available and can be used in different contexts is not new in the logical-philosophical literature[2] and it is also not without problems. In the second part of the chapter, I give this insight a precise formalization, which in the same go solves the associated problems.

2.1 Setting the stage: the *de re*-*de dicto* distinction

The present chapter is about the interaction between propositional attitudes, quantifiers and the notion of identity. All of the conceptual difficulties arising

[1]See Quine (1953), Quine (1956), and Quine (1960). 'What is this object, that denounced Catiline without Philip's having become aware of the fact?' (Quine (1953), p. 147).
[2]See in particular Hintikka (1967) and Hintikka (1969), and more recently Gerbrandy (2000).

from such interplay were first addressed by Quine, in his classic papers in the 50-60s.

Quine (1953) discusses cases of failure of the principle of *substitutivity of identicals.* According to this principle, originally formulated by Leibniz, co-referential expressions are interchangeable everywhere *salva veritate.* Quine considers the following example:

(39) 'Cicero' contains six letters.

Although Cicero is Tully, the substitution of the second for the first does not preserve the truth value of the sentence. The following is false:

(40) 'Tully' contains six letters.

The principle of substitutivity fails to hold in this case.

Another principle which clearly fails in connection with this kind of examples is the principle of *existential generalization.* If we apply such a principle to example (39) we obtain:

(41) $\exists x$('x' contains six letters)

which 'consists merely of a falsehood – namely "The 24th letter of the alphabet contains six letters." – preceded by an irrelevant quantifier.'[3] Contexts like quotations in which substitution of co-referential names may not preserve truth value are called *referentially opaque* by Quine. Given the difficulties illustrated by example (41), Quine has taken the view that quantification into opaque contexts is always misguided.

Propositional attitudes and modalities also create referential opaque contexts. Consider, for instance, belief attributions, as in the examples in (42). The following three sentences are mutually consistent:

(42) a. Philip believes that Cicero denounced Catiline.

b. Cicero is Tully.

c. Philip does not believe that Tully denounced Catiline.

Substitution of co-referential terms in belief contexts is not always allowed. Furthermore, consider sentence (43) which should be derivable from (42a) according to the principle of existential generalization:

(43) $\exists x$(Philip believes that x denounced Catiline)

[3] Quine (1953), p. 147.

The problem with this sentence is that we cannot identify this object that according to Philip denounced Catiline. It cannot be Cicero, that is, Tully because to assume this would conflict with the truth of (42c). Quine concludes that quantification in propositional attitude contexts is *always* unwarranted too, as in the quotation contexts considered above. But consider now the following sentence:

(44) Ralph believes that the president of Russia is bald.

Suppose Ralph believes that Jeltsin is the president of Russia, but, as we all know, Putin is the actual president of Russia. How do we interpret (44)? Does it say that Ralph believes that Jeltsin is bald or Putin? Intuitively (44) can have both readings. On the first reading, the term 'the president of Russia' is presented as part of Ralph's belief and is interpreted from his perspective, so inside the scope of the belief operator. On the second reading, the same description is not taken to belong to Ralph's conceptual repertoire, but it is used to denote the actual president of Russia so that the description is interpreted from our perspective, thus outside the belief operator. Now consider another situation. Suppose Ralph does not have any idea about who the president of Russia is. (44) is again ambiguous. On the first reading, in which 'the president of Russia' is interpreted from Ralph's perspective, Ralph has an unspecific belief about whoever is the president of Russia; on the other reading, in which reference is made from our perspective, it is asserted that Ralph believes of Putin, who is *de facto* the president of Russia, that he is bald.

Now, if we assume the first reading of (44), substitution of co-referential terms can change the truth value of the sentence. Under this reading, (44) and (45) can have different truth values and the principle of substitutivity fails.

(45) Ralph believes that Putin is bald.

On the other hand, substitutivity is warranted, if we assume the second reading. If we interpret the relevant terms from the speaker's perspective, (44) is true iff (45) is true.

Furthermore, while existential generalization can intuitively not be warranted in the first case, it is always in the second case. Consider the second described situation. On the first interpretation of (44), the derivation of (46) is problematic in such a situation, but it is intuitively justified on the second reading.

(46) $\exists x$(Ralph believes that x is bald)

From this example we can conclude that belief contexts are not always opaque. Quine (1953)'s conclusion was too drastic after all. There are readings of belief reports for which the principle of substitutivity of identicals and existential generalization do hold. These have been called *de re* readings. Belief contexts which do not warrant these principles, and so are referentially opaque, are called *de dicto*.

The existence of unproblematic cases of quantification into propositional attitude contexts is recognized by Quine himself in a later paper,[4] where he discusses the following classical example also illustrating the *de re*-*de dicto* contrast:

(47) Ralph believes that someone is a spy.

a. Ralph believes there are spies.

b. There is someone whom Ralph believes to be a spy.

Example (47) is ambiguous between a *de dicto* reading, paraphrased in (47a), asserting that Ralph believes that the set of spies is not empty, and a *de re* reading, paraphrased in (47b), saying that there is a particular individual to whom Ralph attributes espionage. 'The difference is vast: indeed, if Ralph is like most of us, (47a) is true and (47b) is false'.[5]

In the next section, I introduce modal predicate logic and I show how we can deal with *de re* and *de dicto* belief in such a framework.

2.2 Modal Predicate Logic

In this section I introduce possible world semantics for Modal Predicate Logic (MPL). Possible world semantics originates from the work of authors like Carnap, Kanger, Hintikka and Kripke in the late 1950's. Before this date, investigations into modal logic were essentially proof theoretical. In the first part of the chapter, I will just consider modal logic from a model-theoretic perspective. A proof theoretic approach is presented in section 2.4.5, and shown to be sound and complete for the semantics in Appendix A.2.

A *language* $\mathcal{L}$ of modal predicate logic takes as primitive the following symbols:

(1) For each natural number $0 \leq n$ a (possibly finite but at most denumerably infinite) set $\mathcal{P}$ of n-place *predicates*.

(2) A (possibly finite but at most denumerably infinite) set $\mathcal{C}$ of *individual constants*.

(3) A denumerably infinite set $\mathcal{V}$ of *individual variables*.

(4) The symbols $\neg$, $\wedge$, $\exists$, $\Box$, $=$, (, and).

The following formation rules specify which expressions are to count as well formed of our language:

[4] Quine (1956).

[5] Quine (1956), p. 178.

R0 Any individual constant in $\mathcal{C}$ or variable in $\mathcal{V}$ is a term in $\mathcal{L}$.

R1 Any sequence of symbols formed by an n-place predicate followed by n terms is a well formed formula (wff).

R2 If t and t' are terms, $t = t'$ is a wff.

R3 If ϕ is a wff, then $\neg\phi$ is a wff.

R4 If ϕ is a wff and x is a variable, $\exists x\phi$ is a wff.

R5 If ϕ and ψ are wffs, so is $(\phi \wedge \psi)$.

R6 If ϕ is a wff, then $\Box\phi$ is a wff.

R7 Nothing else is a wff.

The standard abbreviations $\phi \rightarrow \psi = \neg(\phi \wedge \neg\psi)$, $\forall x\phi = \neg\exists x\neg\phi$ and $\Diamond\phi = \neg\Box\neg\phi$ are adopted.

A *model* M for $\mathcal{L}_{MPL}$ is a quadruple $\langle W, R, D, I\rangle$ in which W is a non-empty set of possible worlds; R is a relation on W, D is a non-empty set of individuals; and I is an interpretation function which assigns for each $w \in W$ an element $I_w(c)$ of D to each individual constant c in $\mathcal{C}$, and a subset $I_w(P)$ of D^n to each n-ary predicate P in $\mathcal{P}$.

Well-formed expressions in $\mathcal{L}$ are interpreted in models with respect to an assignment function $g \in D^{\mathcal{V}}$ and a world $w \in W$.

2.2.1. Definition. [MPL-Interpretation of Terms]

(i) $[t]_{M,w,g} = g(t)$ if t is a variable.

(ii) $[t]_{M,w,g} = I_w(t)$ if t is a constant.

I define now a satisfaction relation $\models$, holding between a worlds w and a formula ϕ in a model M and relative to an assignment g, saying that ϕ is true in M and w with respect to g.

2.2.2. Definition. [MPL-Interpretation of Formulas]

$$\begin{array}{rcl}
M, w \models_g Pt_1, ...t_n & \text{iff} & \langle [t_1]_{M,w,g}, ..., [t_n]_{M,w,g}\rangle \in I_w(P) \\
M, w \models_g t_1 = t_2 & \text{iff} & [t_1]_{M,w,g} = [t_2]_{M,w,g} \\
M, w \models_g \neg\phi & \text{iff} & \text{not } M, w \models_g \phi \\
M, w \models_g \phi \wedge \psi & \text{iff} & M, w \models_g \phi \text{ and } M, w \models_g \psi \\
M, w \models_g \exists x\phi & \text{iff} & \exists d \in D : M, w \models_{g[x/d]} \phi \\
M, w \models_g \Box\phi & \text{iff} & \forall w' : wRw' : M, w' \models_g \phi
\end{array}$$

A formula is valid in a model M iff it is true with respect to all assignments and all worlds in M. A formula is valid in MPL iff it is valid in all models.

2.2.3. Definition. [MPL-Validity] Let $M = \langle W, R, D, I\rangle$ be a model for $\mathcal{L}$ and ϕ a wff of $\mathcal{L}$.

$$\begin{array}{rcl} M \models_{MPL} \phi & \text{iff} & \forall w \in W, \forall g \in D^{\mathcal{V}} : M, w \models_g \phi \\ \models_{MPL} \phi & \text{iff} & \forall M : M \models_{MPL} \phi \end{array}$$

The idea of using modal predicate logic to represent the logic of propositional attitudes derives from Hintikka (1962). For simplicity, I will deal with cases in which one propositional attitude, namely belief, is attributed to one person only.[6] The set of worlds w accessible from w_0, $\{w \in W \mid w_0 R w\}$, is seen as the belief state $Bel(w_0)$ of the relevant subject in w_0. $Bel(w_0)$ represents the set of the subject's doxastic alternatives in w_0, that is the set of possibilities that are compatible with her belief in that world. A sentence like 'a believes that ϕ' is translated as $\Box\phi$. $\Box\phi$ is true in w_0 iff ϕ is true in all worlds accessible from w_0. This intuitively means that a subject a believes that ϕ is true iff in all possible worlds compatible with what a believes, it is the case that ϕ. a does not believe that ϕ is true iff in at least one world compatible with what a believes, it is not the case that ϕ.[7]

Principles of a Logic of Belief

Kripke (1963) points out that different types of modal notions can be characterized by certain axiom schemes that constraint the accessibility relation. For instance, consider the following scheme which may fail to hold only in a model in which some possibility is inaccessible from itself:[8]

T $\Box\phi \rightarrow \phi$

[6]In order to speak about multiple attitudes $p \in P$ relative to a set of agents A, we can extend the language of predicate logic by adding a set of operators $\Box_a^p$ for each agent $a \in A$ and propositional attitude $p \in P$. A model for such a multi-modal language will be a structure which specifies a set of accessibility relations $(R_a^p)_{a \in A}^{p \in P}$ rather than a single R, where accessibility relations representing different propositional attitudes satisfy different constraints. Of course other interesting and non trivial issues arise once more attitudes and more agents are taken into consideration, such as the representation of common belief (see Fagin et al. (1995) and more recently Gerbrandy (1999)) and the logic of the interactions between operators representing the different attitudes of one agent (see Hintikka (1962) and Heim (1992)). However, since these issues cut across the topic of the present work, they are ignored here.

[7]One objection that has been raised against representing objects of belief in terms of sets of possible worlds is that it fails to account for ignorance of non-contingent matters, such as mathematical or logical truth. Since the problem of logical omniscience is not relevant to the issue I want to discuss in this work, I shall ignore it.

[8]The labels used for the principles in this section are historically motivated. See for instance Hughes and Cresswell (1996) for the relevant references.

T is a plausible principle for certain interpretations of the modal operators, for instance metaphysical necessity or knowledge. If something is necessary true, then it is true. And what is known must be the case. Thus a modal system which wants to capture the logic of these notions should only consider models with a reflexive accessibility relations ($\forall w : wRw$). On the other hand, **T** is not a plausible principle for a logic of belief. If you believe that it is raining, it does not follow that it is raining, because you might be wrong. Belief is not a factive notion, contrary to knowledge. Other schemes, however, are intuitively valid when it concerns beliefs. Belief is normally taken to satisfy positive and negative introspection. If you (do not) believe something, you believe that you (do not) believe it. The following two principles, thus, are plausible for a logic of belief.

4 $\Box\phi \rightarrow \Box\Box\phi$

E $\neg\Box\phi \rightarrow \Box\neg\Box\phi$

4 corresponds to the transitivity of R ($\forall w, w', w'' : wRw' \;\&\; w'Rw'' \;\Rightarrow wRw''$). **E** expresses the fact that R is a euclidean relation ($\forall w, w', w'' : wRw' \;\&\; wRw'' \;\Rightarrow w'Rw''$). A further assumption which is often made (see for instance Hintikka (1962)) is that only consistent belief states are taken into consideration. If you believe ϕ, then ϕ is consistent with your belief state.

D $\Box\phi \rightarrow \Diamond\phi$

This principle is satisfied in all models in which each world has at least one accessible world ($\forall w : \exists w' : wRw'$). A relation which satisfies this condition is called a serial relation.

To summarize, the following three principles, corresponding to the respective conditions on the accessibility relation, will be adopted in what follows:

1. *Consistency* (Serial Relations)

   D $\Box\phi \rightarrow \Diamond\phi$

2. *Positive Introspection* (Transitive Relations)

   4 $\Box\phi \rightarrow \Box\Box\phi$

3. *Negative Introspection* (Euclidian Relations)

   E $\neg\Box\phi \rightarrow \Box\neg\Box\phi$

Reflexivity (Factivity) expressed by principle **T** and Symmetry expressed by the following principle **B**: $\Diamond\Box\phi \rightarrow \phi$ are not assumed. **B** is not a reasonable principle for belief. Not everything which is consistent to believe must be the case.

Belief and Quantification

In this section, we consider how MPL deals with the interaction between quantifiers and the belief operator.

First of all the present semantics validates the following two schemes which are known as the Barcan Formula and its Converse:

BF $\forall x \Box \phi \rightarrow \Box \forall x \phi$

CBF $\Box \forall x \phi \rightarrow \forall x \Box \phi$

Numerous objections have been raised against the intuitive validity of these two principles. The standard way to provide a semantics where the Barcan formula does not hold is to allow for models with increasing domains. In order to do so, a model is defined as a quintuple: $\langle W, R, D, F, I \rangle$ where W, R, D, I are as above, and F is a function from W to subsets of D, which satisfies the following condition: if wRw', then $F(w) \subseteq F(w')$. If we want to falsify the converse of the Barcan Formula, we need to drop the inclusion requirement.[9] However, since considerable difficulties arise if we drop the inclusion requirement, and since the philosophical issue related to the intuitive interpretation of the Barcan Formula and its Converse is not prominent in the present work, I will restrict my discussion to a semantics in which domains are not allowed to vary.

Also the following related 'mixed' principle, which I call the principle of Importation, holds in MPL:

IM $\exists x \Box \phi \rightarrow \Box \exists x \phi$

Its converse, however, which will be called the principle of Exportation, is *not* generally valid:

EX $\Box \exists x \phi \rightarrow \exists x \Box \phi$

The failure of **EX** is crucial for the MPL representation of the *de re*-*de dicto* distinction.

de re and *de dicto*

In MPL, the *de re*-*de dicto* contrast can be expressed by means of permutation of components of formulae.[10] Sentences like (48) or (49) can be assigned the following two logical forms:

[9]See Hughes and Cresswell (1996) for a clear formal discussion of these issues.
[10]See Russell (1905).

(48) Ralph believes that someone is a spy.

a. $\Box\exists x S(x)$

b. $\exists x \Box S(x)$

(49) Ralph believes that the president of Russia is bald.

a. $\Box\exists x(r = x \wedge B(x))$

b. $\exists x(r = x \wedge \Box B(x))$

The (a) logical forms express the *de dicto* readings with possibly different individuals being spies or presidents in different doxastic alternatives of Ralph. The (b) logical forms expresses the *de re* readings, in which one and the same individual is ascribed espionage or baldness in all worlds compatible with what Ralph believes. The *de re-de dicto* distinction is represented in terms of a scope ambiguity. *de re* belief reports are sentences which contain some free variable in the scope of a belief operator. Note that the two logical forms $\Box\exists x(r = x \wedge B(x))$ and $\Box B(r)$ are equivalent in the present semantics. I will use the latter representation for the *de dicto* reading of sentences like (49) in the following discussion.

It is easy to see that by means of these representations, MPL manages to tackle the intuitions about the *de re* and the *de dicto* belief exposed in section 2.1. Let's see first how the referential opacity of *de dicto* belief is accounted for.

MPL invalidates the following unrestricted versions of the principles of substitutivity of identicals[11] and of existential generalization:

SI $t_1 = t_2 \rightarrow (\phi[t_1] \rightarrow \phi[t_2])$

EG $\phi[t] \rightarrow \exists x \phi[x]$

where $\phi[t_1]$ and $\phi[t_2]$ differ only in that the former contains the term t_1 in one or more places where the latter contains t_2. The two principles can fail in the presence of some belief operator when applied to arbitrary singular terms.[12] The reason for this is that the interpretation of a belief operator involves a shift of

[11] The rule of substitution of identicals is more general than it is stated in **SI** and involves expressions of any category. I will just concentrate on the substitutivity of individual terms here.

[12] Substitution of co-referential terms and existential generalization are allowed, if applied to variables or in the absence of any belief operator. The following principles are valid in MPL:

SI1 $t_1 = t_2 \rightarrow (\phi[t_1] \rightarrow \phi[t_2])$ (if ϕ is non-modal)

SIv $x = y \rightarrow (\phi[x] \rightarrow \phi[y])$

EG1 $\phi[t] \rightarrow \exists x \phi[x]$ (if ϕ is non-modal)

EGv $\phi[y] \rightarrow \exists x \phi[x]$

the world of evaluation and two terms can refer to one and the same individual in one world and yet fail to co-refer in some other. Given this fact, we can easily account in MPL for the consistency of the following three sentences, where (50a) and (50c) are assigned a *de dicto* interpretation:

(50) a. Ralph believes that the president of Russia is bald.

$\Box B(r)$

b. The president of Russia is Putin.

$r = p$

c. Ralph does not believe that Putin is bald.

$\neg\Box B(p)$

Even if 'Putin' and 'the president of Russia' denote one and the same man in the 'actual' world, thus making the identity (50b) true, they can refer to different men in the worlds conceived possible by Ralph. For this reason (50a) and (50c) can both be true. The principle of substitutivity of identicals does not hold in general.[13] The substitutivity puzzle involving proper names, illustrated by example (42), can be handled in the same way.

(42) a. Philip believes that Cicero denounced Catiline.

$\Box\phi(c)$

b. Cicero is Tully.

$c = t$

c. Philip does not believe that Tully denounced Catiline.

$\neg\Box\phi(t)$

The intuitive consistency of the three sentences can be accounted for by assuming that in different doxastic alternatives a proper name can denote different individuals. The failure of substitutivity of co-referential terms (in particular proper names) in belief contexts does not depend on the ways in which terms actually refer to objects (so this thesis is not in opposition with Kripke (1972)'s analysis

[13] A weaker version of the principle of substitutivity of identicals holds for sentences containing a belief operator, if we assume consistency, positive and negative introspection:

SI$_\Box$ $\Box t_1 = t_2 \rightarrow (\Box\phi[t_1] \rightarrow \Box\phi[t_2])$

If we are discussing what a person believes we can substitute a term for another iff they refer to one and the same individual in all her doxastic alternatives. If we consider all models rather than only serial, transitive, and euclidean models, the principle holds only if ϕ is non-modal.

of proper names), it is simply due to the possibility that two terms that actually refer to one and the same individual are not believed by someone to do so.[14]

A similar line of reasoning explains why we cannot always existentially quantify with respect to a term which occurs in a belief context. The term 'the president of Russia' may be such that it refers to different men in Ralph's doxastic alternatives (and in the actual world), and therefore the *de dicto* reading (51a) of a sentence like (51) does not necessarily imply (52) or (53):

(51) Ralph believes that the president of Russia is bald.

a. $\Box B(r)$

b. $\exists x(x = r \wedge \Box B(x))$

(52) There is someone whom Ralph believes to be bald.

$\exists x \Box B(x)$

(53) Ralph believes Putin to be bald.

$\exists x(x = p \wedge \Box B(x))$

On the other hand, both (52) and (53) are obviously derivable[15] from the *de re* reading (51b) of (51) and, therefore, the referential transparency of *de re* belief is also accounted for.

The contrast between the *de re* and the *de dicto* logical forms is a genuine one. Existential generalization and exportation of terms from belief contexts is not generally allowed. In order for existential generalization (or term exportation) to be applicable to a term t occurring in the scope of a belief operator, t has to denote the same individual in all doxastic alternatives of the relevant agent (plus the actual world). The following two principles are valid in MPL, if we assume consistency, positive and negative introspection:[16]

$\mathbf{EG}_\Box$ $\exists x \Box x = t \rightarrow \Box\phi[t] \rightarrow \exists x \Box \phi[x]$

$\mathbf{TEX}_\Box$ $\exists x(x = t \wedge \Box x = t) \rightarrow \Box\phi[t] \rightarrow \exists x(x = t \wedge \Box\phi[x])$

Sentences like $\exists x(x = t \wedge \Box x = t)$ are used by a number of authors, notably Hintikka, as representations of knowing-who constructions. $\mathrm{TEX}_\Box$ says that a term t is exportable from a belief context if we have as an additional premise that the relevant subject knows who t is. In MPL, having a *de re* belief implies knowing who somebody is.

[14] See chapter 1, footnote 10.

[15] Example (52) follows by simple reasoning. Example (53) is derived by substitution of co-referential terms which holds if operated outside the scope of a belief operator (with the assumption that Putin is the president of Russia).

[16] If we consider also non-serial, non-transitive and non-euclidean models, the two principles are valid only if ϕ does not contain any belief operator.

The double vision puzzles

Although MPL can account for the *de re-de dicto* distinction, its solution to the substitutivity paradox is not fully satisfactory. Since variables refer to one and the same individual in all possible worlds, the following version of the substitutivity principle holds in MPL:

SIv $x = y \rightarrow (\phi[x] \rightarrow \psi[y])$

As argued by Church (1982), substitutivity paradoxes can be constructed which depend on variables rather than descriptions or names. It is easy to see from **SIv** that MPL validates the following scheme:

LIv $x = y \rightarrow \Box x = y$

Furthermore in all serial frames, the following is valid as well:[17]

CLNIv $\Box x \neq y \rightarrow x \neq y$

Now consider the formulation of the two principles in 'quasi-ordinary' language:

(54) For every x and y, if $x = y$, then George IV believes that $x = y$.

(55) For every x and y, if George IV believes that $x \neq y$, then $x \neq y$.

Example (54) can be understood to say that one individual cannot be believed by George IV to be two, for instance, George IV cannot fail to recognize as the same individual, an individual encountered on two different occasions. Example (55) says that George IV can make individuals distinct by merely believing that they are distinct. In MPL, George IV, as well as anyone with consistent beliefs, is predicted to have such incredible powers. These two predictions can intuitively be accepted 'only on the doubtful assumption that belief properly applies "to the fulfillment of condition by objects" quite "apart from special ways of specifying" the objects'.[18] Following Church, I call this assumption the principle of *transparency of belief.* In the literature, a series of so called double vision situations

[17] MPL also validates the following scheme:

LNIv $x \neq y \rightarrow \Box x \neq y$

And in all serial models also the following is valid:

CLIv $\Box x = y \rightarrow x = y$

LNIv and **CLIv** are not discussed by Church, since they are not derivable by simple application of substitutivity.

[18] This is a quote from Church (1982), p. 62, who quotes Quine (1953), p. 151.

have been discussed which illustrate the problematic nature of such a principle. In all of these examples, we find someone who knows an individual in different guises, without realizing that it is one and the same individual.

A famous case is the one discussed by Quine (1956). Quine tells of a man called Ralph, who ascribes contradictory properties to Ortcutt since, having met him on two quite different occasions, he is 'acquainted' with him in two different ways. Another well-known example is described in Kripke (1979). In Kripke's story, the bilingual Pierre assents to 'Londres est jolie' and denies 'London is pretty', because he does not recognize that the ugly city where he lives now, which he calls 'London', is the same city as the one he calls 'Londres', about which he has heard while he was in France. In a third situation, described in Richards (1993), a man does not realize that the woman to whom he is speaking through the phone is the same woman he sees across the street and who he perceives to be in some danger. In such a situation the man might sincerely utter: 'I believe that she is in danger', but not 'I believe that you are in danger', although the two pronouns 'she' and 'you' refer to one and the same woman. Let me expand upon the situation discussed in Quine (1956):

> There is a certain man in a brown hat whom Ralph has glimpsed several times under questionable circumstances on which we need not enter here; suffice it to say that Ralph suspects he is a spy. Also there is a grey-haired man, vaguely known to Ralph as rather a pillar of the community, whom Ralph is not aware of having seen except once at the beach. Now Ralph does not know it but the men are one and the same.[19]

Consider the following sentence:

(56) Ralph believes that x is a spy.

$\Box S(x)$

Is (56) true under an assignment which maps x to the individual Ortcutt which is the man seen on the beach *and* the man with the brown hat?[20] As Quine observes, the ordinary notion of belief seems to require that although (56) holds when x is specified in one way, namely as the man with the brown hat, it may yet fail when the same x is specified in some other way, namely as the man seen on the beach. Belief 'does not properly apply to the fulfillment of conditions by *objects* apart from special ways of specifying them'[21]. In standard modal predicate logic, we cannot account for this ordinary sense of belief. Since variables range over bare individuals, we cannot account for the fact that sentences like (56) depend on the way of specifying these individuals. This feature also implies that in MPL

[19] Quine (1956), p. 179.

[20] This question is structurally identical to the initial question of this chapter.

[21] See Quine (1953), p. 151.

the following two sentences cannot both be true, unless we want to charge Ralph with contradictory beliefs:

(57) Ralph believes Ortcutt to be a spy.

$\exists x(x = o \wedge \Box S(x))$

(58) Ralph believes Ortcutt not to be a spy.

$\exists x(x = o \wedge \Box \neg S(x))$

In MPL we cannot avoid the inference from (57) and (58) to (59):

(59) Ralph believes Ortcutt to be a spy and not to be a spy.

$\exists x(x = o \wedge \Box(S(x) \wedge \neg S(x)))$

Example (59) says that Ralph's belief state is contradictory. This prediction clashes with our intuitions about the Ortcutt case. On the one hand, since Ralph would assent to the sentence: 'That man with the brown hat is a spy', we are intuitively allowed to infer (57). On the other hand, since 'Ralph is ready enough to say, in all sincerity, 'Bernard J. Ortcutt is no spy',' [22] we are also ready to infer (58). But this should not imply that Ralph has contradictory beliefs. Ralph is not 'logically insane', he simply lacks certain information. In the following section, I will discuss three refinements of the standard modal predicate semantics which have been proposed as a response to Quine's intriguing puzzle, and I will show that they also raise problems of their own.

2.3 Contingent Identity Systems

Consider the property P which an object x has iff Ralph believes of x that x is a spy. From the discussion in the previous section, it seems that it depends on the way of referring to x whether P applies to x. A number of systems have been proposed that try to account for this dependence. Here I just consider what Hughes and Cresswell (1996) call Contingent Identity (CI) systems, based on the framework of modal predicate logic. In CI systems, the principles of necessary (non-)identity **LIv** and **LNIv** and their converse do not hold. This result is obtained by allowing a variable to take different values in different worlds. A standard way to do this is to let variables range over so-called individual concepts.[23] As in the previous chapter, an *individual concept* is a total function from possible worlds in W to individuals in D.

[22] Quine (1956), p. 179.

[23] This use of individual concepts can be seen to be anticipated in Frege (1892), and Carnap (1947).

2.3.1 Quantifying over all Concepts

In the first Contingent Identity semantics we will consider (CIA), variables are taken to range over *all* individual concepts in $IC = D^W$. Language and models are defined as in MPL. Well-formed expressions are interpreted in models with respect to a world and an assignment function $g \in IC^{\mathcal{V}}$. Variables are crucially assigned concepts in IC, rather than individuals in D. The denotation of a variable x with respect to an assignment g and a world w is the instantiation $g(x)(w)$ of $g(x)$ in w.

2.3.1. DEFINITION. [CI-Interpretation of variables] $[x]_{M,w,g} = g(x)(w)$

In the semantics, we only have to adjust the clause dealing with existential quantification.

2.3.2. DEFINITION. [Quantification over all individual concepts]

$$M, w \models_g \exists x\phi \quad \text{iff} \quad \exists c \in IC : M, w \models_{g[x/c]} \phi$$

It is easy to see that CIA does not validate **LIv** and **LNIv** and their converses. Let $M = \langle W, R, D, I\rangle$ be such that $W = \{w, w'\}$ and $R = \{\langle w, w'\rangle, \langle w', w'\rangle\}$; and let $g, g' \in IC^{\mathcal{V}}$ be such that $g(x)(w) = g(y)(w)$ and $g(x)(w') \neq g(y)(w')$, and $g'(x)(w) \neq g'(y)(w)$ and $g'(x)(w') = g'(y)(w')$ respectively. Then we have:

(i) $M, w, g \not\models_{CIA} x = y \rightarrow \Box x = y$;

(ii) $M, w, g \not\models_{CIA} \Box x \neq y \rightarrow x \neq y$.

(iii) $M, w, g' \not\models_{CIA} x \neq y \rightarrow \Box x \neq y$

(iv) $M, w, g' \not\models_{CIA} \Box x = y \rightarrow x = y$

By invalidating **LIv**, CIA avoids the double vision paradoxes. Given the situation described by Quine, a sentence like (56) is true under an assignment which maps x to the concept $\lambda w[\text{the man with the brown hat}]_w$ and false under an assignment which maps x to the concept $\lambda w[\text{the man seen on the beach}]_w$.

(56) Ralph believes that x is a spy.
$\Box S(x)$

In this way the dependency of belief on the ways of specifying the intended objects is accounted for. Furthermore, (57) and (58) below do not entail the problematic (59):

(57) $\exists x(x = o \wedge \Box S(x))$

(58) $\exists x(x = o \wedge \Box\neg S(x))$

(59) $\exists x(x = o \wedge \Box(S(x) \wedge \neg S(x)))$

We can infer the following, but it does not entail that Ralph's beliefs are inconsistent:

(60) $\exists x(x = o \wedge \exists y(o = y \wedge \Box(S(x) \wedge \neg S(y))))$

The analysis of *de re* belief reports formalized by CIA can be intuitively formulated as follows:

> The sentence 'a believes b to be ϕ' is true iff there is a representation α such that α is actually b and a believes that α is ϕ.

To believe *de re* of x that x has P is to ascribe P to x under some representation.

Although in a different framework, Quine (1956)[24] predicts similar truth conditions for *de re* belief attributions. In order to account for ordinary cases of quantification into belief contexts, Quine distinguishes two notions of belief, *notional* (belief$_1$) and *relational* (belief$_2$). The latter contains one or more of the crucial terms in a purely referential position and therefore sustains both substitution of identicals and existential generalization. For instance, a sentence like:

(61) Ralph believes that Ortcutt is a spy.

is assigned two interpretations, represented as follows:

(62) Ralph believes$_1$ ('Ortcutt is a spy')

(63) Ralph believes$_2$('x is a spy', Ortcutt)

According to Quine, the relational interpretation (63), which corresponds to the *de re* reading, is implied[25] by any sentence of the form:

(64) Ralph believes$_1$ ('α is a spy')

together with the simple identity 'α = Ortcutt'.

As noticed by Kaplan (1969), there is a problem with this analysis. Upon a closer inspection, Quine's account (or CIA more in general) fails to capture the intuitions that originally led us to a distinction between the *de re* and *de dicto* representations. The shortest spy problem illustrates why.

[24] Quine considers intensions (individual concepts but also propositions) 'creatures of darkness' (Quine (1956), p. 180) and analyzes his notional belief reports as relations between individuals and sentences rather than propositions. I will disregard this issue here.

[25] 'The kind of exportation which leads from (62) to (63) should doubtless be viewed in general as implicative.' Quine (1956), p. 182.

The shortest spy problem

It is easy to see that CIA validates the principle of exportation:

EX $\Box\exists x\phi \rightarrow \exists x\Box\phi$

It follows that the general form of existential generalization **EG** and term exportation from belief contexts are also validated:

TEX $\Box\phi[a] \rightarrow \exists x(x = a \wedge \Box\phi[x])$

The following two examples illustrate why the validity of these schemes clashes with our intuitions about *de re* belief.

Consider the following case discussed in Kaplan (1969). Suppose Ralph believes there are spies, but does not believe of anyone in particular that she is a spy. He further believes that no two spies have the same height which entails that there is a shortest spy. In such a situation, the *de re* reading of Quine's spy example (65), which 'was originally intended to express a fact that would interest the F.B.I.'[26], is intuitively false.

(65) There is someone whom Ralph believes to be a spy.

The problem of the present semantics is that such a reading is not captured by the following representation:

(66) $\exists x\Box S(x)$

Given the circumstances described above, we have no troubles in finding a value for x under which $\Box S(x)$ is true, namely the concept $\lambda w[\text{the shortest spy}]_w$, therefore, (66) is true in CIA and hence cannot be used to express (65). The classical representation of the *de re*-*de dicto* contrast in terms of scope permutation is no longer available. The following example illustrates the same difficulty. Consider the *de re* sentence:

(67) Ralph believes Putin to be the president of Russia.

(67) is intuitively false in a situation in which Ralph believes that Jeltsin is the president. In CIA, however, the standard *de re* representation of (67), namely (68) is implied by any sentence of the form (69) together with the simple identity $\alpha = p$.

(68) $\exists x(x = p \wedge \Box x = r)$

[26] Kaplan (1969), p. 220.

(69) Ralph believes that α is the president of Russia.

$\Box\, \alpha = r$

In particular, it is implied by a trivially true sentence like:

(70) Ralph believes that the president of Russia is the president of Russia.

$\Box\, r = r$

in case Putin is *de facto* the president of Russia. Thus, in CIA, (68) is true in the described situation and therefore cannot serve as a representation (67).

The validity of (term) exportation conflicts with our intuitions about the significant difference between the *de re* and *de dicto* readings of belief attributions. One conclusion we could draw from these examples is simply the inadequacy of an analysis of *de re* belief which assumes that they involve quantification over concepts, rather than objects. **EX** and **TEX** are indeed very natural principles, if we take quantifiers to range over concepts. If a believes that someone is the so and so, then there is a concept, viz. the so and so, such that a believes that it is the so and so. On the other hand, if we assumed, instead, that *de re* belief applies to bare individuals rather than concepts, we would go back to MPL with its double vision difficulties. Many authors have therefore maintained that an analysis of *de re* belief involving quantification over ways of specifying individuals is on the right track. What is needed, if we want to solve the difficulties presented in the present section, is not a return to quantification over individuals rather than representations, but 'a frankly inequalitarian attitude'[27] towards these representations. This is Quine's diagnosis of the 'shortest spy cases', further developed by Kaplan (1969).[28] According to this view, *de re* belief attributions do involve quantifications over representations, yet not over all representations. 'The shortest spy' or 'the president of Russia' in the examples above are typical instances of representations that should not be allowed in our domain of quantification. In the next subsection, I present and investigate a second contingent identity system which formally works out this strategy. It will turn out, however, that also this kind of analysis is not fully satisfactory.

2.3.2 Quantifying over Suitable Concepts

Standard MPL was too stringent in allowing only plain individuals as possible values for our variables, and CIA was too liberal in allowing all concepts to count as 'objects'. An adequate semantics might be one which allows some (possibly non-rigid) concepts to count as possible values for our variables, but not all.

[27]The quotation is from the end of Quine (1961).

[28]'... a solution might lie in somehow picking out certain kind of names as being required for the exportation.' Kaplan (1969), p. 221.

Such a semantics could be obtained by taking models which specify which sets of concepts are to count as domains of quantification. Such a *model* will be a quintuple $\langle W, R, D, S, I\rangle$ in which W, R, D, I are as above and $S \subseteq IC$. In this second Contingent Identity semantics CIB, variables are taken to range over a subset of the set of all concepts. Assignments $g \in S^{\mathcal{V}}$ map individual variables to elements of S.

2.3.3. Definition. [Quantification over suitable concepts]

$$M, w \models_g \exists x\phi \quad \text{iff} \quad \exists c \in S : M, w \models_{g[x/c]} \phi$$

It is easy to see that the notion of validity defined by CIB is weaker than the notion of MPL and of CIA validity. Indeed, all MPL and CIA models are CIB models, properly understood. CIB-validity thus entails MPL- and CIA-validity, but not the other way around.

2.3.4. Proposition. Let ϕ be a wff of $\mathcal{L}$.

(i) $\models_{CIB} \phi \;\Rightarrow\; \models_{MPL} \phi$

(ii) $\models_{CIB} \phi \;\Rightarrow\; \models_{CIA} \phi$

proof: the proof relies on the fact that for each MPL or CIA model $M = \langle W, R, D, I\rangle$ we can build two corresponding CIB models, M_{MPL} and M_{CIA}, which satisfy conditions (a) and (b) respectively, for all wffs ϕ:

(a) $M_{MPL} \models_{CIB} \phi \;\Leftrightarrow\; M \models_{MPL} \phi$

(b) $M_{CIA} \models_{CIB} \phi \;\Leftrightarrow\; M \models_{CIA} \phi$

The construction of the two models is straightforward. Let $M = \langle W, R, D, I\rangle$ be an MPL (or CIA) model. We build $M_{MPL} = \langle W', R', D', S_{MPL}, I'\rangle$ and $M_{CIA} = \langle W', R', D', S_{CIA}, I'\rangle$ as follows. We let W', R', D', I' be like W, R, D, I and S_{MPL} and S_{CIA} contain all and only rigid concepts and all concepts respectively: $S_{MPL} = \{\lambda w[d] \mid d \in D\}$ and $S_{CIA} = IC$. It is an easy exercise to see that conditions (a) and (b) are satisfied. But then for S ranging over MPL and CIA, $\not\models_S \phi$ implies for some M, $M \not\models_S \phi$ which implies for the corresponding model M_S, $M_S \not\models_{CIB} \phi$, which means $\not\models_{CIB} \phi$. □

Note that MPL-validity does not entail CIA-validity or the other way around:

(iii) $\models_{MPL} \phi \;\not\Rightarrow\; \models_{CIA} \phi$

(iv) $\models_{CIA} \phi \;\not\Rightarrow\; \models_{MPL} \phi$

Together with proposition 2.3.4, this implies that CIB-validity is strictly weaker than MPL- and CIA-validity.

(v) $\models_{MPL} \phi \quad \not\Rightarrow \quad \models_{CIB} \quad \phi$

(vi) $\models_{CIA} \phi \quad \not\Rightarrow \quad \models_{CIB} \phi$

The CIB semantics is very promising. By proposition 2.3.4, clause (i), CIB does not validate **EX** or **TEX**, since they are not valid in MPL. Therefore, it seems to avoid the shortest spy problems. On the other hand, by proposition 2.3.4, clause (ii), **LIv** is also invalidated, since it fails to hold in CIA, and therefore the double vision difficulties do not arise.

In CIB, *de re* belief attributions are analyzed as follows:

> The sentence 'a believes b to be ϕ' is true iff there is a *suitable* representation α such that α is actually b and a believes that α is ϕ.

To believe *de re* of x that x has P is to ascribe P to x under some suitable representation α.

Although it uses a different framework,[29] the influential analysis in Kaplan (1969) can be classified in this group. In that article, Kaplan attempts a concrete characterization of the notion of a suitable representation with respect to an agent and an object. A necessary and sufficient condition for the truth of a sentence of the form 'a believes x to be P', is the existence of a representation α in the conceptual repertoire of the agent a such that (i) α denotes x, (ii) α is a name *of* x for a, (iii) α is sufficiently vivid, and (iv) a believes α is P.

So, for instance,

(68) Ralph believes Putin to be the president of Russia.

$\exists x(x = p \wedge \Box x = r)$

is accepted iff there is a vivid name α of Putin in Ralph's conceptual repertoire such that Ralph believes that α is the president of Russia.

I will not discuss Kaplan's analysis in detail, but just note that *de re* belief reports are analyzed as describing mental acts or states of the agent. Their truth or falsity depends on a fact about the belief state as such, and this is in accordance with a kind of semantics like CIB in which the set of suitable representations is selected by the model, rather than by a contextual parameter, since the model also fully determines the belief state of the one relevant subject.[30]

[29] Kaplan (1969) follows Quine (1956) in assuming that objects of belief are sentences and not propositions, but by using Frege's method of representation of intermediate contexts he manages to account for the *de re-de dicto* ambiguity by permutation of scope rather than by positing two primitive senses of beliefs.

[30] If we consider multi-modal extensions of the semantics, different subjects can be taken to have different conceptual repertoires. So in order to properly extend Kaplan's analysis to these models, we should take different sets of suitable concepts S_a as assigned to different agents $a \in A$.

In CIB, the problematic exportation steps in the shortest spy and in the president examples are blocked simply by assuming that S does not contain the concepts $\lambda w[\text{the shortest spy}]_w$ or $\lambda w[\text{the president of Russia}]_w$, since the two descriptions are clearly not vivid names of the intended objects for our Ralph in the described circumstances. However, the examples below show that there still are problems with this type of analysis.

Odette's lover and other problems

In CIB, the existence of a suitable representation α of b for a such that a believes that α is ϕ is a *necessary* and a *sufficient* condition for the truth of the *de re* sentence 'a believes b to be ϕ'. I call this condition condition (A). The examples I will discuss here show that this biimplication leads to a number of empirical difficulties. In the case of Odette's lover below (and that of Susan's mother), we find a counterexample to the necessity of condition (A). The theater case shows condition (A) not to be sufficient either. This double failure finds a natural explanation once we recognize the context dependence of the notion of a suitable representation. On different occasions, different sets of representations can count as suitable depending on the circumstances of the conversation, rather than on the mental state of the relevant agent. The problem of CIB is that models encode the information about what are the suitable representations and, therefore, they are not equipped to account for this variability.

Odette's lover Consider the following situation described by Andrea Bonomi.

> Thanks to some clues, Swann has come to the conclusion that his wife Odette has a lover, but he has no idea who his rival is, although some positive proof has convinced him that this person is going to leave Paris with Odette. So he decides to kill his wife's lover, and he confides his plan to his best friend, Theo. In particular, he tells Theo that the killing will take place the following day, since he knows that Odette has a rendezvous with her lover. [...] Unknown to Swann, Odette's lover is Forcheville, the chief of the army, and Theo is a member of the security staff which must protect Forcheville. During a meeting of this staff to draw up a list of all the persons to keep under surveillance, Theo (who, unlike Swann, knows all the relevant details of the story) says:
>
> (71) Swann wants to kill the chief of the army.
>
> meaning by this that Swann is to be included in the list. The head of the security staff accepts Theo's advice. [...] Swann is kept under surveillance. A murder is avoided.[31]

[31] Bonomi (1995), pp. 167-168.

Let's see whether CIB can account for this example. On its *de dicto* reading, (71) is false for obvious reasons. On its *de re* reading, it is true only if we can find a relevant α among the suitable representations or, in Kaplan's terminology, among the vivid names for Swann of Forcheville, such that Swann wants to kill α. The only possible candidate in the described situation is the description 'Odette's lover'. Formally, the sentence is true if the concept $\lambda w[\text{Odette's lover}]_w$ is in S. Now according to the most intuitive characterization of the notion of a vivid name of x for a, the relevant concept should not count as a suitable one. Swann does not know who Odette's lover is. Condition (A) is not satisfied, and, therefore, CIB cannot account for the truth of Theo's report. In order to avoid this counterintuitive result a proponent of CIB could argue in favor of a weaker notion of a suitable representation. However, if our notion of a suitable representation were as weak as to tackle this example, it would be too weak to solve the shortest spy problem. If in order to account for the truth of (71) we say that 'Odette's lover' is a vivid name of Forcheville for Swann, then the following sentences would also be true in the described situation:

(72) Swann believes of the chief of the army that he is Odette's lover.

which seems to be unacceptable, intuitively. To summarize, either 'Odette's lover' counts as a suitable description of Forchevilles for Swann, or it does not. If it does not, condition (A) is not satisfied for either example and so CIB fails to account for the truth of (71). If it does, condition (A) is satisfied for both examples and, therefore, CIB fails to account for the unacceptability of (72). A natural way out of this impasse would be to accept that one and the same representation can be suitable in one occasion and not in another. But if the set of suitable descriptions is determined by the model as in CIB, this solution is not available.

In the following example, due to van Fraassen, we find another case illustrating the same point.

Susan's mother

> Susan's mother is a successful artist. Susan goes to college, where she discusses with the registrar the impact of the raise in tuition on her personal finances. She reports to her mother 'He said that I should ask for a larger allowance from home'. Susan's mother exclaims: 'He must think I am rich!' Susan, looking puzzled, says 'I don't think he has any idea who you are'.[32]

van Fraassen analyzes the example as follows:

> The information the mother intends to convey is that the registrar believes that Susan's mother is rich, while Susan misunderstands her

[32] van Fraassen (1979), pp. 371-372.

> as saying that the registrar thinks that such and such successful artist is rich. The misunderstanding disappears if the mother gives information about herself, that is, about what she had in mind. She relied, it seems, on the auxiliary assertion 'I am your mother'.[33]

I repeat the crucial sentence uttered by Susan's mother:

(73) He must think I am rich.

The utterance of the mother may be not fully felicitous, because it is ambiguous, but it is not false; indeed, Susan can accept the sentence after the clarification of her mother. Again a proponent of CIB faces a dilemma: either (a) he does not accept 'Susan's mother' as a vivid representation of the referent of the pronoun 'I' in (73) for the registrar or (b) he does. If (a), then CIB fails to account for the intuitive acceptability of (73); if (b), it fails to account for the unacceptability of the following sentence when uttered by Susan's mother in the same situation:

(74) He must think I am your mother.

Again, CIB has difficulties in explaining the difference in acceptability between the two utterances, because it has no explanation of why one and the same representation 'Susan's mother' can be used on one occasion and not on the other.

What the examples of Odette's lover and of Susan's mother show is that the cognitive relation between the agent (Swann or the registrar) and the object of belief (Forcheville or the mother) does not always play an essential role in deciding about the acceptability of the *de re* sentences. In both examples, a term t is exported also if the agent does not have an intuitively acceptable answer to the question 'Who is t?'. Whether a representation is suitable or not depends in the two cases on what properties are ascribed under such a representation, that is, on a fact about the conversation rather than on Swann's or the registrar's belief state. An approach upon which the information about which concepts are suitable is stored in the model fails to account for such dependencies.[34]

In the following case, we find a counterexample to the sufficiency of condition (A) and a further illustration of the evident context sensitivity of *de re* constructions.

[33] van Fraassen (1979), p. 372.

[34] That the context of utterance (in particular the intentions of the participant in the conversation) is relevant for the interpretation of *de re* belief attribution more than the belief state of the agent itself has been observed among others by van Fraassen (1979), Stalnaker (1988) and Crimmins and Perry (1989), and more recently, again, in van Rooy (1997) and Gerbrandy (2000).

the theater Consider the following situation described in the 1999 edition of Bonomi (1983). Suppose Leo correctly believes that Ugo is the only one in town who has a bush jacket. Leo further believes that Ugo has climbed the Cervino mountain. One night Ugo lends his bush jacket to another friend of Leo, Pio. Wearing it, Pio goes to the theater. Leo sees him and believes he is Ugo and utters:

(75) That person has climbed the Cervino.

From then on, nothing happens to modify Leo's belief about the climbing abilities of his friends. By the way, he is informed of the well-known fact that Pio hates climbing and he would never say that Pio climbed the Cervino. Now consider sentence (76) uttered by a fourth person Teo in two different circumstances α and β.

(α) Two months later.

(β) The same evening at the theater immediately after Leo's utterance of (75).

(76) Leo believes that Pio has climbed the Cervino.

Bonomi observes a contrast between the acceptability of the sentence in the two contexts. In α, the sentence is hardly acceptable, although, as it seems, condition (A) is satisfied,[35] and so constitutes a counterexample to the sufficiency of such condition.

On the other hand, our intuitions about the acceptability of (76) in context β are less sharp and the sentence might be acceptable if uttered immediately after Leo manifestation of his beliefs. In the following enriched context γ, the acceptability of (76) may be more evident:

(γ) Leo after uttering (75) goes to Pio and congratulates him for his great performance. Giò, a fourth friend, asks Teo for an explanation of Leo's surprising behaviour.

It seems to me that in γ, (76) can be accepted as an appropriate answer to Giò's question. All of the considered analyses have difficulties in explaining this kind of context relativity. For all of them, the *de re* reading of (76) is an eternal proposition whose truth value depends on Leo's belief state, which has not changed with

[35]'Indeed, Leo has a 'name' α of Pio (namely 'the man wearing a bush jacket at the theater') such that: (i) α actually denotes Pio; (ii) α is in a causal relation with Pio (since it is originated in a perceptive contact); (ii) and α is a sufficiently vivid name, for Leo, of Pio (again because of the perceptive contact). In addition, Leo believes that the man wearing a bush jacket at the theater has climbed the Cervino mountain.' (Bonomi (1999), my translation.) Therefore, at least if we assume Kaplan's characterization of the notion of a suitable representation, condition (A) is satisfied in this case.

respect to the relevant facts. So, sentences like (76) are either true or false, irrespective of the context in which they are uttered. CIA would predict that (76) is true since we have a description under which Leo believes of Pio that he climbed the mountain. A proponent of MPL or CIB would have to decide whether the perceptive contact between Pio and Leo at the theater counts as an acquaintance relation. If it does, then we have a suitable representation of Pio for Leo under which Leo makes the relevant attribution and this is a sufficient condition for the truth of (76) also in context α. If it does not, then by the necessity of condition (A) we fail to account for the appropriateness of (76) in (γ).

As a last example, consider the following variation on Quine's celebrated double vision puzzle, which constitutes a further illustration of the context relativity of *de re* belief reports.

Ortcutt again You can tell each half of the Ortcutt story separately. In one half Ralph sees Ortcutt wearing the brown hat. In the other he sees him on the beach. From the first story you can reason as in (77). From the second story as in (78).

(77) a. Ralph believes that the man with the brown hat is a spy.

b. The man with the brown hat is Ortcutt.

c. So Ralph believes of Ortcutt that he is a spy.

(78) a. Ralph believes that the man seen on the beach is not a spy.

b. The man seen on the beach is Ortcutt.

c. So Ralph does *not* believe of Ortcutt that he is a spy.

Although we don't have to assume that there is any change in Ralph's belief state, it seems unproblematic to say that Ralph believes Ortcutt to be a spy and Ralph does *not* believe Ortcutt to be a spy, depending on which part of the story you are taking into consideration. The challenge is how to account for the compatibility of (77c) and (78c). Their natural representations (79) and (80) respectively are obviously contradictory:

(79) $\exists x(x = o \wedge \Box S(x))$

(80) $\neg\exists x(x = o \wedge \Box S(x))$

Proponents of CIB (or CIA) could then argue that a correct representation for (78c) is not (80), but rather (81) which, in CIB, is not in contradiction with (79):

(81) $\exists x(x = o \wedge \neg\Box S(x))$

Contrary to MPL, in which the two logical forms (80) and (81) are equivalent, CIA and CIB, predict a structural ambiguity for sentences like ‘Ralph does *not* believe Ortcutt to be a spy’, with a wide scope reading asserting that Ralph does not ascribe espionage to Ortcutt under any (suitable) representation, and a narrow scope reading asserting that there is a (suitable) representation under which Ralph does not ascribe espionage to Ortcutt. This ambiguity is automatically generated by any system which assumes that *de re* belief reports involve quantification over representations of objects rather than over the objects themselves. Intuitively, however, it is hard to detect an ambiguity in the natural language sentences.[36] Furthermore, such an account of the possible consistency of (77c) and (78c), would lack an explanation of the influence of the previous discourse on the acceptability of one *or* the other sentence. Intuitively (77c) is acceptable in (77), but not in (78) because the relevant description, namely ‘the man with the brown hat’, which is explicitly mentioned in (77), is absent in (78) and not salient in that context. Again CIB fails to account for this type of context sensitivity.

To summarize, in the first two cases we have seen how one and the same description (‘Odette’s lover’ and ‘Susan’s mother’) can or cannot be a suitable representation of an intended object for an agent whose beliefs are described, this depending on the circumstances of the utterance and the property ascribed. In the last two cases, we have seen one and the same *de re* belief report (examples (76) and (77c)) which obtains different truth values when it is uttered in different circumstances without any relevant change in the belief state of the subject. CIB, which assumes that *de re* belief attributions involve quantification over a set of suitable concepts determined by the model of interpretation, cannot account for any of these cases without automatically generating other problems.

From the examples discussed in this section we can conclude that although the problem of interpreting quantification into the scope of a belief operator can be seen as the problem of distinguishing suitable representations from non-suitable ones, it is not the cognitive relation between the subject of belief and the intended object alone, that can supply the central notion for this distinction. Rather it seems that other elements play a crucial role, namely the conversational circumstances in which the belief report is made, the property ascribed, and the interests and goals of the participants in the conversation. The pragmatic analysis in the following section tries to take into account such dependencies. In section 2.4 it is worked out more systematically.

[36] Such an ambiguity does not seem to have been empirically observed before the above-mentioned philosophical theory has been proposed. It seems an unexpected consequence of the theory rather than a meant prediction. Indeed, as far as I know, nobody has ever argued in favor of it.

2.3.3 Pragmatic Analysis

In the previous section, I discussed a number of examples illustrating the dependence of *de re* belief on various pragmatic factors. Those examples suggest the following preliminary rough analysis of *de re* belief reports:

> The sentence 'a believes b to be ϕ' uttered in context C is true iff there is a description α *suitable in* C such that α is actually b and a believes that α is ϕ.

Formally, a model is defined as in standard MPL and assignments are defined as in CIA. The interpretation function is relativized to a pragmatic parameter which selects sets of contextually salient concepts out of D^W. Let Z be a set of concepts whose value is pragmatically supplied.

2.3.5. Definition. [Quantification over contextually selected concepts]

$$M, w, Z \models_g \exists x\phi \quad \text{iff} \quad \exists c \in Z : M, w, Z \models_{g[x/c]} \phi$$

The idea behind the pragmatic analysis is that *de re* belief reports express different contents in different contexts, in the same way (or in a similar way) as sentences containing indexical expressions. In different circumstances, different sets of concepts are assumed to supply the domain of quantification of our quantifiers. The interpretation of *de re* belief reports, which directly depends on how objects are identified across the boundary of different possible worlds, is crucially affected by this variability.

The analysis I propose in the next section is among the pragmatic approaches. It has however the following extra feature that I believe is not trivial and is not a matter of detail. It is assumed that not all sets of concepts can be pragmatically selected as domains of quantification, but only those satisfying two natural conditions. The first condition is that for each individual d in the domain and each world w, the selected set Z must contain a concept which identifies d in w. The second condition is that Z cannot contain overlapping concepts, i.e. concepts standing for one and the same individual in one world and for two different individuals in another. I call the former the *existence* condition and the latter the *uniqueness* condition. In what follows I will present some arguments in favor of their assumption.

The question whether an individual can fail to be identifiable in Z in some world (*existence*) is equivalent to the question whether existential generalization can fail, if applied to wffs which do not contain any belief operator:

EG1 $\phi[t] \rightarrow \exists x\phi[x]$ (if ϕ is non-modal)

We expect principle **EG1** to hold in our semantics. Contrast the following two examples:[37]

[37] Assume that the president of Russia is not a character of fiction.

(82) a. If Ralph believes that the president of Russia is a spy, then there is someone whom Ralph believes to be a spy.

b. $\Box S(r) \to \exists x \Box S(x)$

(83) a. If the president of Russia is a spy, then there is someone who is a spy.

b. $S(r) \to \exists x S(x)$

While, as we have seen, (82) can intuitively fail to be generally valid, (83) cannot. The failure of existential generalization is a peculiarity of opaque contexts. Existential generalization intuitively holds in the absence of belief operators. Assuming the existence condition for our quantificational domains is a natural way of accounting for this intuition. Note that CIB, which, as I presented it, does not assume such condition, does not validate **EG1**.[38]

The question whether we should allow overlapping concepts (*uniqueness*) is equivalent to the question whether the following principles can fail to be true assuming that x and y range over one and the same quantificational domain Z:

(84) $\forall x \forall y (x = y \to \Box x = y)$

(85) $\forall x \forall y (x \neq y \to \Box x \neq y)$

Intuitively, if x and y stand for individuals we expect the principles to hold, if they stand for representations of individuals we expect them to fail. Individuals do not split when we move from one world to the other, whereas, a characteristic property of representations is that two representations can coincide in one situation (denote one and the same individual) and split or not in another. Now, recall the principles of exportation **EX**: $\Box \exists x \phi \to \exists x \Box \phi$ and of term exportation **TEX**: $\Box \phi[a] \to \exists x (x = a \wedge \Box \phi[x])$. As we have already seen, if we take variables to range over representations, rather than objects, then **EX** or **TEX** are intuitively plausible, but then instances of their conclusions cannot be used to express the *de re* reading of natural language belief reports. It seems fair to conclude that this is a clear sign that *de re* readings involve quantification over genuine objects, rather than over ways of specifying them. By adopting principles like (84) and (85), and so the uniqueness condition, we can capture this intuition. The following example supplies further empirical justification for the uniqueness condition.

Consider again the Ortcutt story. Further assume that Ortcutt and Portcutt are the only two members of the local anti-X club and Ralph believes Portcutt to be a spy. Now it seems to me that in such a situation the following sentences can be true:

[38] In order to avoid this counterintuitive result, a proponent of CIB can either assume that the set of suitable concepts S satisfies the existence condition or it can relativize the interpretation function I to S. The latter strategy is obviously not available in a pragmatic approach, where the choice of I is prior to the selection of Z.

(86) Of each member of the anti-X club, Ralph believes that he is a spy.
$\forall x(X(x) \rightarrow \Box S(x))$

A possible reaction to (86) could be the following:

(A) I don't accept (86), because I don't accept one of the following two:

(87) Ralph believes Ortcutt to be a spy.
$\exists x(x = o \wedge \Box S(x))$

(88) Ralph believes Portcutt to be a spy.
$\exists x(x = p \wedge \Box S(x))$

But it would be rather weird to react as in (B):

(B) I accept (87) and (88), but I don't accept (86), because Ralph does not ascribe espionage to Ortcutt under all relevant representations, for instance he does not under the representation 'the man seen on the beach'.

Reaction (B) would be the reaction of someone quantifying over a set containing the overlapping concepts 'the man with the brown hat' and 'the man seen on the beach'. Such a reaction is rather out of place. Intuitively we do not accept reply (B), because we expect the universal quantifier in (86) to quantify over the objects Ortcutt and Porcutt and not over their representations, which are essentially overlapping. By adopting the uniqueness condition we can account for these intuitions.

There is one last point which we need to clarify. Consider again Quine's double vision situation. We expect our semantics to account for the fact that the following sentences are mutually consistent in that situation and do not imply that Ralph has inconsistent beliefs:

(89) Ralph believes Ortcutt to be a spy.
$\exists y(y = o \wedge \Box S(y))$

(90) Ralph believes Ortcutt not to be a spy.
$\exists x(x = o \wedge \neg\Box S(x))$

The CI solution to this puzzle crucially involves the presence in the quantificational domain of two overlapping concepts, namely *the man on the beach* and *the man with the brown hat* which stand for one individual in one world and for two different individuals in another. If we rule out overlapping concepts altogether it is not immediately clear how we can account for this case. On the one hand, *de re* belief reports are about individuals, as I have argued above. On the other, their interpretation crucially depends on the ways of specifying these individuals, as illustrated by the Ortcutt case. The problem is how to combine these two intuitions. MPL accounts for the first intuition, but fails to capture the double vision cases. The CI systems' solution of the double vision puzzles, leads directly to the problems discussed in this section. A pragmatic approach supplies us with a natural way out from this impasse. The compatibility of the two sentences (89) and (90) is captured by letting the variables x and y range over different sets of concepts. The availability of different sets of non-overlapping concepts as possible domains of quantification on different occasions, enables us to account for the dependence of belief reports on the ways of referring to objects (and so for double vision cases), without dropping the uniqueness condition, and so avoiding the counterintuitive results described in this section.

The conclusion that can be drawn from this discussion is that *de re* belief reports neither involve bare quantification over individuals *simpliciter* (MPL), nor over ways of specifying these individuals (CI), but rather over the individuals themselves, only specified in one determinate way.

2.4 Conceptual Covers in Modal Predicate Logic

In this section, I present Modal Predicate Logic under Conceptual Covers (CC). In section 2.4.1, I restate the definition of the notion of a conceptual cover introduced in the previous chapter. Section 2.4.2 presents the semantics of CC. Section 2.4.3 discusses a number of applications. Section 2.4.4 compares the CC notion of validity with the classical MPL one. Finally, section 2.4.5 introduces an axiom system which provides a sound and complete characterization of the set of CC-valid wffs.

2.4.1 Conceptual Covers

A conceptual cover is a set of individual concepts that satisfies the following condition: in a conceptual cover, in each world, each individual constitutes the instantiation of one and only one concept.

Given a set of possible worlds W and a universe of individuals D, a *conceptual cover* CC based on (W, D) is a set of functions $W \to D$ such that:

$$\forall w \in W: \ \forall d \in D: \ \exists! c \in CC: c(w) = d$$

Conceptual covers are sets of concepts which exhaustively and exclusively cover the domain of individuals. In a conceptual cover each individual d is identified by at least one concept in each world (*existence*), but in no world is an individual counted more than once (*uniqueness*).

It is easy to prove that each conceptual cover and the domain of individuals have the same cardinality.[39] In a conceptual cover, each individual is identified by one and only one concept. Different covers constitute different ways of conceiving one and the same domain.

Illustration Consider the following situation. In front of Ralph stand two women. For some reason we don't need to investigate, Ralph believes that the woman on the left, who is smiling, is Bea and the woman on the right, who is frowning, is Ann. As a matter of fact, exactly the opposite is the case. Bea is frowning on the right and Ann is smiling on the left. In order to formalize this situation, we just need to distinguish two possibilities. The simple MPL model $\langle W, R, D, I\rangle$ visualized by the following diagram will suffice:

$$\begin{array}{cccc} w_1 & \mapsto & (\ddot{\smile}) & (\ddot{\frown}) \\ & & [ann] & [bea] \\ \\ w_2 & \mapsto & (\ddot{\smile}) & (\ddot{\frown}) \\ & & [bea] & [ann] \end{array}$$

W consists of two worlds w_1 and w_2. w_2 is the only world accessible from w_1 for Ralph. D consists of two individuals $(\ddot{\smile})$ and $(\ddot{\frown})$. As illustrated in the diagram, in w_1, which stands for the actual world, Ann is the woman on the left, whereas in w_2, which represents the one possibility in Ralph's doxastic state, Bea is the woman on the left.

There are only two possible conceptual covers definable over such sets of worlds W and individuals D, namely:

$$\text{A} = \{\lambda w[\text{left}]_w, \lambda w[\text{right}]_w\}$$

$$\text{B} = \{\lambda w[\text{Ann}]_w, \lambda w[\text{Bea}]_w\}$$

These two covers corresponds to the two ways of cross-identifying individuals (i.e. telling of an element of a possible world whether or not it is identical with a given element of another possible world) which are available in such a situation: A cross-identifies those individual which stand in the same perceptual relation to Ralph. B cross-identifies the women by their name.

All other possible combinations of concepts fail to satisfy the existential or the uniqueness condition. For instance the set C is not a conceptual cover:

$$\text{C} = \{\lambda w[\text{left}]_w, \lambda w[\text{Ann}]_w\}$$

[39] See chapter 4, proposition 4.1.2.

Formally, C violates both the existential condition (no concept identifies (☹) in w_1) and the uniqueness condition ((☺) is counted twice in w_1). Intuitively, the inadequacy of C does not depend on the individual properties of its two elements, but on their combination. Although the two concepts *the woman on the left* and *Ann* can both be salient, when contrasted with each other, they cannot be regarded as standing for the two women in the universe of discourse in all relevant worlds.

When we talk about concepts, we implicitly assume two different levels of 'objects': the individuals (in D) and the ways of referring to these individuals (in D^W). An essential feature of the intuitive relation between the two levels of the individuals and of their representations is that to one element of the first set correspond many elements of the second. The intuition behind it is that one individual can be identified in many different ways. What characterizes a set of representations of a certain domain is this cardinality mismatch, which expresses the possibility of considering an individual under different perspectives which may coincide in one world and not coincide in another. Individuals, on the other hand, do not split or merge once we move from one world to the other. Now, since the elements of a cover also cannot merge or split (by uniqueness), they behave like individuals in this sense, rather than representations. On the other hand, a cover is not barely a set of individuals, but encodes information on how these individuals are specified. We thus can think of covers as sets of individuals each identified in one specific way. My proposal is that *de re* belief reports involve quantification over precisely this kind of sets. By allowing different conceptual covers to constitute the domain of quantification in different occasions, we can account for the double vision cases, without failing to account for the intuition that *de re* belief reports involve quantification over genuine individuals, rather than over ways of specifying these individuals.

2.4.2 Quantification under Cover

A *language* of modal predicate logic under conceptual covers $\mathcal{L}_{CC}$ is the language formed out of $\mathcal{L}$ by the addition of a set of new primitive symbols N of conceptual cover indices $0, 1, 2, ...$ and by changing the definitions of the rules R0 and R4 as follows:

R0′ (i) If α is a variable in $\mathcal{V}$ and n is a CC-index, α_n is a term.

(ii) If α is an individual constant in $\mathcal{C}$, α is a term.

R4′ If ϕ is a wff, x_n is an indexed variable, then $\exists x_n \phi$ is a wff.

CC-indices range over (contextually selected) conceptual covers. I will write $\mathcal{V}_n$ to denote the set of variables indexed with n and $\mathcal{V}_N$ to denote the set $\bigcup_{n \in N}(V_n)$.

A *model* for $\mathcal{L}_{CC}$ is a quintuple $\langle W, R, D, I, C\rangle$ in which W, R, D, I are as above and C is a set of conceptual covers over (W, D).

2.4.1. Definition. [CC-Assignment] Let $K = C^N \cup IC^{\mathcal{V}_N}$. A CC-assignment g is an element of K satisfying the following condition: $\forall n \in N$: $\forall x_n \in \mathcal{V}_n$: $g(x_n) \in g(n)$.

A CC-assignment g has a double role, it works on CC-indices and on indexed variables. CC-indices are assigned to conceptual covers elements of C and n-indexed individual variables x_n are assigned to concepts elements of $g(n)$.

The definition of quantification is relativized to conceptual covers. Quantifiers range over elements of contextually determined conceptualizations.

2.4.2. Definition. [Quantification under Cover]

$$M, w \models_g \exists x_n \phi \quad \text{iff} \quad \exists c \in g(n) : M, w \models_{g[x_n/c]} \phi$$

All other semantic clauses are defined as in MPL, as well as the notion of validity.

Illustration Consider again the situation described above, with two women standing in front of Ralph. Imagine now that Bea is insane and Ralph is informed about it, Ann is, instead, vaguely known to Ralph as a quiet school teacher. He still wrongly thinks that Bea is the woman on the left and Ann is the woman on the right. We can formalize the situation by the following CC model $\langle W, R, D, I', C \rangle$, in which W, R, D are as above, I' is like I with the only addition that Bea is insane in w_1 and in w_2 (in the diagram insanity is represented by a bullet), and C contains the two covers A and B introduced above:

$$\begin{array}{llll} w_1 & \mapsto & (\ddot\smile) & (\ddot\frown)^\bullet \\ & & [ann] & [bea] \\ & & & \\ w_2 & \mapsto & (\ddot\smile)^\bullet & (\ddot\frown) \\ & & [bea] & [ann] \end{array}$$

Consider now the following *de re* sentence:

(91) Ralph believes Ann to be insane.

$\exists x_n(x_n = a \wedge \Box I(x_n))$

(91) will have different contents when interpreted under different conceptual covers. Under an assignment which maps n to cover A, i.e., if the operative conceptual cover is the one which cross-identifies objects by pointing at them, the sentence is true. As a matter of fact, Ann is the woman on the left and Ralph ascribes insanity to the woman on the left.

On the other hand, if the operative cover is the one which cross-identifies objects by their name, than (91) is false, Ralph indeed believes that Bea is insane, and not Ann.

This variability is in accordance with our intuitions. The acceptability of the sentence is relative to the circumstances of the utterance. For instance, (91) could be correctly uttered as an explanation of Ralph's weird behaviour, when all of a sudden he starts chasing the woman on the left to bring her to a mental institution. But it might be harder to accept for instance as an answer to a question about Ralph's general beliefs about Ann.

2.4.3 Applications

The theoretical point behind the present analysis is that natural language *de re* belief reports are about individuals under a perspective. The uniqueness and the existential conditions on conceptual covers exemplify the idea that in *de re* belief reports we never explicitly quantify over ways of specifying objects, but over the objects themselves. On the other hand, the dependency of *de re* belief on the ways of specifying the intended objects is accounted for by allowing different sets of concepts to count as domains of quantification on different occasions. The former feature helps in avoiding the shortest spy problem, the latter provides a solution to the double vision puzzles.

double vision and the theater

Since variables can range over elements of different conceptual covers, the present semantics does not validate the principles **LI** and **LNI** of 'necessary' (non)-identity or their converses. In CC, we can express cases of mistaken identity. The sentences (92) and (93) can be true also in serial models, if the indices n and m are assigned two different covers.

(92) $\exists x_n \exists y_m (x_n \neq y_m \wedge \Box x_n = y_m)$

(93) $\exists x_n \exists y_m (x_n = y_m \wedge \Box x_n \neq y_m)$

The consistency of sentences like (93) show that we can deal with double vision situations. Recall Quine's Ralph who believes of one man Ortcutt that he is two distinct individuals, because he has seen him in two different circumstances, once with a brown hat, once on the beach. If we want to represent this sort of situation we have to use two different conceptual covers. Representations of cases of mistaken identity crucially involve shifts of conceptualization. This reflects the fact that, intuitively, in order to describe Ralph's misconception, the speaker must assume two different ways of identifying the objects in the domain. According to one way, Ortcutt is identified as the man with the brown hat, according to the other as the man seen on the beach. Shifts of covers are expensive. We will see in chapter 4 that they are acceptable by the audience only in order to repair violations of general pragmatic constraints and require reasoning and adjustments. The necessity of a plurality of conceptualizations in order to represent double vision cases explains the extraordinary nature of such situations.

In order to see how the specific examples in Quine's story are accounted for in the present framework, consider the very simple model $M = \langle W, R, D, C, I\rangle$. W consists of only two worlds, the actual world w_0 and w_1. R is such that w_1 is the only world accessible from w_0. D consists of two individuals Ortcutt o and Porcutt p. In w_0, Ortcutt is the man with the brown hat, but he is also the man seen on the beach. In w_1, Ortcutt is the man on the beach and Porcutt is the man with the brown hat. In both w_0 and w_1, p is a spy and o is not. M can be used to model the Ortcutt situation, by assuming that $\{w_1\} = Bel(w_0)$ represents Ralph's belief state. There are four concepts definable in such a model:

$$\begin{array}{ccc} & & \mathbf{a} \\ \hline w_0 & \mapsto & o \\ w_1 & \mapsto & o \end{array} \qquad \begin{array}{ccc} & & \mathbf{b} \\ \hline w_0 & \mapsto & p \\ w_1 & \mapsto & p \end{array} \qquad \begin{array}{ccc} & & \mathbf{c} \\ \hline w_0 & \mapsto & o \\ w_1 & \mapsto & p \end{array} \qquad \begin{array}{ccc} & & \mathbf{d} \\ \hline w_0 & \mapsto & p \\ w_1 & \mapsto & o \end{array}$$

Concept **a** is the interpretation in M of the description 'the man on the beach'. Concept **c** is the interpretation of the description 'the man with the brown hat'.

With respect to W, the concepts **a** and **c** cannot be elements of one and the same cover, because they overlap in w_0 and split in w_1. In order to express Ralph's mistake we have to use two different covers.

$$\begin{array}{cccc} & & \mathbf{a} & \mathbf{b} \\ \hline w_0 & \mapsto & o & p \\ w_1 & \mapsto & o & p \end{array} \qquad\qquad \begin{array}{cccc} & & \mathbf{c} & \mathbf{d} \\ \hline w_0 & \mapsto & o & p \\ w_1 & \mapsto & p & o \end{array}$$

Since we have dropped the assumption that variables within belief contexts refer to bare individuals, we can now give a reasonable answer to Quine's question:

> Can we say of this man (Bernard J. Ortcutt to give him a name) that Ralph believes him to be a spy?[40]

namely, it depends. The question receives a negative or a positive answer relative to the way in which Ortcutt is specified. In the model described above, (94) is true under the assignment that maps x_n to **c** (representing 'the man with the brown hat') and false under the assignment that maps x_n to **a** (which stands for 'the man on the beach'):

(94) $\Box S(x_n)$

As a consequence of this, the following two sentences are true under an assignment which maps n to the cover $\{\mathbf{c}, \mathbf{d}\}$ and m to $\{\mathbf{a}, \mathbf{b}\}$ (we assume that the constant o refers to the individual Ortcutt in w_0):

[40] Quine (1956), p. 179.

(95) Ralph believes Ortcutt to be a spy.
$\exists x_n(x_n = o \wedge \Box S(x_n))$

(96) Ralph believes Ortcutt not to be a spy.
$\exists x_m(x_m = o \wedge \Box\neg S(x_m))$

(95) and (96) can both be true even in a serial model, but only if n and m are assigned different conceptual covers. This is reasonable, because intuitively one can accept these two sentences without drawing the conclusion that Ralph's beliefs are inconsistent, only if one takes into consideration the two different perspectives under which Ortcutt can be considered. On the other hand, the fact that a shift of cover is required in this case explains the never ending puzzling effect of the Ortcutt story. After reading Quine's description of the facts, both covers (the one identifying Ortcutt as the man with the brown hat, the other identifying Ortcutt as the man on the beach) are equally salient, and this causes bewilderment in the reader who has to choose one of the two in order to interpret each *de re* sentence.

From (95) and (96) we *cannot* conclude the following (for $i \in \{n, m\}$):

(97) $\exists x_i(x_i = o \wedge \Box(S(x_i) \wedge \neg S(x_i)))$

which would charge Ralph with contradictory beliefs. Yet, we can conclude (98) which does not carry such a charge:

(98) $\exists x_n(x_n = o \wedge \exists y_m(o = y_m \wedge \Box(S(x_n) \wedge \neg S(y_m))))$

Consider now what happens to the concepts **a**, i.e. 'the man on the beach' and **c**, 'the man with the brown hat', if restricted to Ralph's belief state $Bel(w_0) = \{w_1\}$:

$$\begin{array}{lc|cc} & & \mathbf{a} & \mathbf{c} \\ \hline w_1 & \mapsto & o & p \end{array}$$

If we restrict our attention to Ralph's doxastic alternatives the two concepts do constitute a conceptual cover, since they exhaust the domain and there is no overlap. But, as we have seen, as soon as we take w_0 into consideration, the set $\{a, c\}$ is no longer a conceptual cover:

$$\begin{array}{ll|cc} & & \mathbf{a} & \mathbf{c} \\ \hline w_0 & \mapsto & o & o \\ w_1 & \mapsto & o & p \end{array}$$

The number of definable covers is relative to the number of possible worlds under consideration.[41] A set of concepts that overlap or split with respect to a class

[41] See chapter 4, proposition 4.1.3.

of possible worlds, may cease to do this – and so constitute a conceptual cover – with respect to a smaller class of possibilities. Ralph is in a state of maximal (though incorrect) information, in which all covers coincide.[42] In such a state, all possible contrasting methods of cross-identifying the individuals in the universe collapse into one, according to which the man seen on the beach and the man with the brown hat are just two different objects. Indeed, Ralph may reason as follows from his perspective:

(99) The man on the beach is tall. The man with the brown hat is tall. So all relevant people are tall.

From our perspective, however, such a reasoning is flawed. Once we take the actual world w_0 into consideration, we know that the two descriptions are different ways of specifying one and the same object. The concepts **a** and **c** cannot be elements of one and the same cover and hence we cannot quantify over them.

Other double vision puzzles are treated in a similar way, in particular, Kripke's case of Pierre. Recall Pierre is a bilingual who assents to 'Londres est jolie' while denying 'London is pretty', because he is ignorant about the fact that London and Londres are one and the same town. The two relevant sentences are represented as follows:

(100) a. Pierre believes that London is pretty.

$$\exists x_n(x_n = l \wedge \Box P(x_n))$$

b. Pierre believes that London is not pretty.

$$\exists x_m(x_m = l \wedge \neg P(x_m))$$

(100a) can be true only in a *de re* interpretation[43] in which the relevant singular term is interpreted from the speaker's point of view, who knows that London is Londres, rather than from Pierre's perspective who does not know precisely that. Pierre indeed ascribes 'ugliness' to London under the representation 'London', so the *de dicto* reading would be false. But since there is an actual representation of London, namely 'Londres' under which Pierre ascribes it the opposite property, the sentence can be true if interpreted *de re* under the right conceptualization.

[42] See chapter 4, corollary 4.1.4.

[43] Note that Kripke says that the two sentences are intuitively true in their *de dicto* interpretation and still should not imply that Pierre's beliefs are inconsistent. However, given our intuitive characterization of *de dicto* belief this does not seem correct. Indeed, we could say that (100a) results from an application of **SI** from the sentence:

(101) Pierre believes that Londres is pretty.

where 'London' and 'Londres' are co-referential terms belonging to different languages. But *de dicto* belief does not seem to allow **SI** even if the two co-referential terms are part of different languages.

If we assume that two different covers are operative in the two cases, we account for the intuitive truth of the *de re* sentences (100a) and (100b) without ascribing Pierre contradictory beliefs.

By means of the same mechanism we can account for the context dependence of *de re* sentences illustrated by the theater and Ortcutt cases in section 2.3.2. Since different covers can be assigned to different occurrences of quantifiers, the general principle of renaming is invalidated in the present semantics. The following scheme is *not* valid in CC (where $\phi[x_n/y_m]$ denotes the result of substituting the variable y_m for the variable x_n in the wff ϕ):

PR $\exists x_n \phi \rightarrow \exists x_m \phi[x_n/x_m]$

The relevant counterexamples are sentences in which ϕ contains some belief operator:

(102) $\exists x_n \Box P x_n \not\rightarrow \exists x_m \Box P x_m$

If two different ways of conceptualizing the domain are operative, that is, if two different notions of what counts as a determinate object are assumed, we have no guarantee that if there is a determinate object (according to one cover) such that the subject believes that she is P, then there is a determinate object (according to the other cover) such that the subject believes that she is P.

The failure of **PR** allows us to account for the theater and Ortcutt cases. Let me expand upon the latter. Recall the following examples in which each half of the Ortcutt story is told separately:

(103) Ralph believes that the man with the brown hat is a spy.
The man with the brown hat is Ortcutt.
So Ralph believes of Ortcutt that he is a spy.

(104) Ralph believes that the man seen on the beach is not a spy.
The man seen on the beach is Ortcutt.
So Ralph does *not* believe of Ortcutt that he is a spy.

Consider now the two relevant *de re* sentences:

(105) $\exists x_n (x_n = o \wedge \Box S(x_n))$

(106) $\neg \exists x_m (x_m = o \wedge \Box S(x_m))$

The two sentences do not contradict each other, if n and m are assigned two different covers. And since covers are pragmatically chosen, two different covers can be selected in the two circumstances. The story in (103) in which the concept 'the man with the brown hat' has been explicitly introduced strongly suggests a cover containing this representation (e.g. cover $\{c, d\}$ in our model). Whereas a cover containing 'the man seen on the beach' (e.g. cover $\{a, b\}$ in our model) is made salient by the previous discourse in the second case (104). Finally note that (106) and the following:

(107) $\exists x_m(x_m = o \wedge \neg\Box S(x_m))$

turn out to be equivalent in the present framework, like in classical MPL. This distinguishes the present semantics from the two CI semantics discussed above, which, as we saw, predicted a structural ambiguity for sentences like 'Ralph does *not* believe Ortcutt to be a spy'.[44] On the present account, negations of *de re* belief reports are not structurally ambiguous, but, like their positive counterparts, they are simply context dependent.

the shortest spy and Odette's lover

The present semantics obviously does not validate **EX** or **TEX**. Existential generalization (and term exportation) can be applied to a term t occurring in the scope of a belief operator only with the extra premise that there is a c in the operative cover such that a denotes instantiations of c in all doxastic alternatives of the relevant agent (plus the actual world). This is what the following two CC-valid principles say:[45]

$$\mathbf{EG}_{\Box n} \quad \exists x_n \Box t = x_n \rightarrow (\Box\phi[t] \rightarrow \exists x_n \Box\phi[x_n])$$

$$\mathbf{TEX}_{\Box n} \quad \exists x_n(x_n = t \wedge \Box x_n = t) \rightarrow (\Box\phi[t] \rightarrow \exists x_n(x_n = t \wedge \Box\phi[x_n]))$$

As in MPL, a term occurring in a belief context must denote one and the same determinate object in all of the relevant worlds in order for existential generalization or term exportation to be applicable to it. But unlike in MPL, the notion of a determinate object is not left unanalyzed. What counts as an object is not given *a priori*, but depends on the operative cover, which is contextually determined. The parallelism between standard modal predicate logic and the present semantics with respect to (term) exportation is sufficient to solve the shortest spy problem at least to a certain extent:

(108) a. Ralph believes that there are spies.

$\Box\exists x_n S(x_n)$

b. There is someone whom Ralph believes to be a spy.

$\exists x_n \Box S(x_n)$

[44] Note that it was the presence of overlapping concepts, which express the possibility of considering an object simultaneously under different perspectives, which gave rise to the dubious ambiguity in the CI systems. If each object is identified by one and only one concept, the two readings collapse, as we intuitively expect.

[45] Again we must assume consistency, positive and negative introspection. If we consider also non-serial, non-transitive and non-euclidean CC-models, the two principles are valid only if ϕ does not contain any belief operator.

(109) a. Ralph believes that the president of Russia is the president of Russia.
$\Box r = r$

b. Ralph believes Putin to be the president of Russia.
$\exists x_m(x_m = p \wedge \Box x_m = r)$

As in MPL, (108a) and (109a) express *de dicto* readings with possibly different determinate objects – according to the operative cover – being spies or presidents in different worlds in Ralph's belief state; whereas (108b) and (109b) express *de re* readings, in which one and the same determinate object – according to the operative cover – is ascribed the relevant property in all relevant worlds. Example (108b) does not follow from (108a), and (109b) does not follow from (109a) (plus the assumption $p = r$), because the relevant conceptual covers do not have to include problematic concepts like *the shortest spy* or *the president of Russia* respectively. Still, our semantics allows statements like (108b) and (109b) to be true in circumstances in which intuitively they are obviously deviant, like in the situations described in the shortest spy section above. Even in such situations, we have no trouble in finding values for n and for m under which (108b) and (109b) are accepted, namely any two covers containing the concept λw[the shortest spy]$_w$ or λw[the president of Russia]$_w$ respectively. It is not immediately obvious how we can rule out such problematic assignments. Ruling them out by not including the problematic covers in C in our model is not a viable option, because of the problems of Odette's lover or of Susan's mother discussed above. Recall the relevant sentences in the case of Susan's mother:

(110) He must think I am rich.

(111) He must think I am your mother.

If in order to explain the inadequacy of (111) in the described situation we rule out any cover containing the concept 'Susan's mother', we are unable to account for the truth of (110) in which such concept is crucially quantified over. Hintikka (1962) (or MPL) and Kaplan (1969) (or CIB) cannot account for these cases. In CC, on the other hand, they find a natural explanation. Since covers are contextually selected, this pragmatic procedure will have to satisfy general pragmatic constraints. Interpreting (111) under a cover containing the concept λw[Susan's mother]$_w$ would make the sentences trivially true, and in ordinary circumstances this leads to a violation of general rules of conversation.[46] On the other hand, the assignment of the same cover for the interpretation of (110) would not involve such a violation. The same kind of pragmatic explanation can be employed in order to deal with the counter-intuitive interpretation of examples like (108b) or (109b) discussed above. Also in these two cases, the problematic assignments are ruled out because they cause violations of general pragmatic constraints. In chapter 4, more will be said about these examples.

[46] Grice's Quantity Maxim: Be as informative as is required. See chapter 4 for more discussion.

de re attitudes and knowing-who constructions

As a last application, I wish to briefly discuss the relation between *de re* attitudes and knowing-who constructions. On Hintikka's (1962) and Kaplan's (1969) accounts, having a *de re* attitude requires knowing who somebody is. This seems correct in most of the cases, but not all. In section 2.3.2, we have discussed two examples, notably the cases of Odette's lover and that of Susan's mother, in which having a *de re* attitude did not seem to require knowing who somebody is under any intuitive interpretation of the latter notion. Let's see how this issue is dealt with in the present analysis.

As we have already seen, the following version of the principle of exportation is valid in CC if we assume consistency, positive and negative introspection:

TEX$_{\Box n}$ $\exists x_n(x_n = t \wedge \Box x_n = t) \rightarrow (\Box\phi[t] \rightarrow \exists x_n(x_n = t \wedge \Box\phi[x_n]))$

TEX$_{\Box n}$ can be paraphrased as follows: a term a is exportable *under a specific conceptual cover* if the relevant subject knows who a is *under the same cover*. In CC, we can account for ordinary cases in which having a *de re* attitude requires knowing who somebody is once we recognize that there is no absolute notion of 'knowing who'.[47] As an illustration, consider the following situation. Suppose Ralph has no idea who the actual president of Russia is. But, for some reason, he believes that the president of Russia is bald and nobody else in Russia is bald. Consider the following *de re* sentence:

(112) Ralph believes Putin to be bald.

$\exists x_n(x_n = p \wedge \Box B(x_n))$

Intuitively the sentence is false and the present analysis can easily explain why. The only way we can derive (112) in the described situation, is by exportation of the description 'the president of Russia' from the *de dicto* sentence (113) (and by substituting 'Putin' for the exported description):

(113) Ralph believes that the president of Russia is bald.

$\Box B(r)$

The sentence I used to set the context: 'Suppose Ralph has no idea who the actual president of Russia is' suggests that a conceptualization is prominent under which the following sentence is clearly not satisfied:

(114) Ralph knows who the president of Russia is.

$\exists x_n(x_n = r \wedge \Box\, x_n = r)$

[47] As I showed in chapter 1, knowing who-constructions show the same context dependence as *de re* attitude attributions, and their analysis requires the same machinery we are employing here, their interpretation being relative to the operative method of cross-identification.

Since (114) is false under the suggested conceptual cover, say CC, we cannot export the relevant term from (113) and, therefore, (112) is false under CC. Since we are reading a thesis and not a report of the F.B.I., we have no reason to think that all the examples in it are true, so we just stick to the suggested conceptualization and reject (112). In ordinary situations, in which we do not shift conceptualization, *de re* belief requires knowing who somebody is.

On the other hand, the CC analysis can also explain the cases in which such a requirement is not satisfied. These are cases in which a *de re* sentence is not satisfied under the prominent conceptualization, but it is still accepted as correct for one or other reason. Typical examples of these situations are the cases of Odette's lover or Susan's mother. As an illustration, consider again the latter case. Recall the crucial sentence:

(115) He must think I am rich.

$$\exists x_n(x_n = I \wedge \Box R(x_n))$$

In the described situation, (115) must be obtained by exportation of the description 'Susan's mother' from (116) (and then by substituting 'I' for the exported description):

(116) The registrar must think that Susan's mother is rich.

On the other hand, from Susan's remark: 'I don't think he knows who you are', we understand that the prominent cover in the described situation, say CC, is one which falsifies sentence (117):

(117) The registrar knows who Susan's mother is.

$$\exists x_n(x_n = m \wedge \Box\, x_n = m)$$

This means that 'Susan's mother' is not exportable under CC and, therefore, (115) is false under such a cover. But, intuitively, (115) was acceptable in the described situation. In the present framework, we can account for this, by assuming that for some pragmatic reason, a *shift* of conceptualization is triggered in such a situation.[48] The index n can be mapped to a cover CC^* containing the concept 'Susan's mother', under which both (115) and (117) are true, and so (115) can be accepted. But still we are not ready to accept (117). How do we explain this? Although (117) must be *true* under some conceptualization, otherwise (115) is not acceptable, this does not imply that (117) must be also *acceptable* under that conceptualization. As we have already seen, if interpreted under CC^*, (117) is

[48] See chapter 4, where I argue that the sentence is acceptable in the described situation first of all by charity (we expect Susan's mother to say the truth), but also because (115) is among the best candidates Susan's mother could have chosen in order to express what she wanted to express on that occasion.

trivialized and, therefore, pragmatically incorrect. Having a *de re* belief does not always require knowing who somebody is in a non-trivial way.

As in the double vision cases, the extraordinary nature of these examples shows from the fact that their interpretation involves a shift of conceptualization. That we can account for these cases in our logic is mirrored by the fact that the principle of renaming **PR** does not generally hold. Example (117) must be true under CC^*, but need not be true under any other cover which would not make its interpretation trivial.

(118) $\exists x_n(x_n = m \wedge \Box\, x_n = m) \not\rightarrow \exists x_m(x_m = m \wedge \Box\, x_m = m)$

2.4.4 MPL and CC validity

In this section, I compare CC with ordinary MPL. I will show that, if there are no shifts of conceptual covers, modal predicate logic under conceptual covers is just ordinary modal predicate logic, the two types of semantics turn out to define exactly the same notion of validity. Once we allow shifts of covers though, a number of problematic MPL-valid principles cease to hold.

I call a model for $\mathcal{L}_{CC}$ containing a single conceptual cover a classical model:

2.4.3. Definition. [Classical CC-Models] Let $M = \langle W, R, D, I, C\rangle$ be a model for a language $\mathcal{L}_{CC}$ of modal predicate logic under conceptual covers. M is *classical* iff $|C| = 1$.

I define a notion of classical CC-validity. A formula is classically valid iff it is valid in all classical CC-models.

2.4.4. Definition. [Classic CC-Validity] Let ϕ be a wff in $\mathcal{L}_{CC}$.

$$\models_{CCC} \phi \quad \text{iff} \quad \forall M : M \text{ is classical } \Rightarrow M \models_{CC} \phi$$

If we just consider classical models, the logic of conceptual covers does not add anything to ordinary modal predicate logic. Classical CC-validity is just ordinary MPL-validity.

The main result of this section is expressed by the following proposition where ϕ is a wff in $\mathcal{L}_{CC}$ which is clearly also interpretable in modal predicate logic:[49]

2.4.5. Proposition. Let ϕ be a wff in $\mathcal{L}_{CC}$.

$$\models_{CCC} \phi \quad \text{iff} \quad \models_{MPL} \phi$$

[49]Hughes and Cresswell (1996), pp. 354-356 show a similar result, namely that Lewis's counterpart theory and Modal Predicate Logic define the same notion of validity if the counterpart relation C is assumed to satisfy the following conditions: (a) C is an equivalence relation; and (b) an individual has one and only one counterpart in each world. In chapter 4 we will see that conceptual covers and counterpart relations which satisfy these two conditions flesh out exactly the same notion.

One direction of the proof of this proposition follows from the fact that given a classical CC-model M, we can define an equivalent ordinary modal predicate logic model M', that is, an MPL-model that satisfies the same wffs as M. Let M be $\langle W, R, D, I, \{CC\}\rangle$. We define an equivalent model $M' = \langle W', R', D', I'\rangle$ as follows. $W' = W$, $R' = R$, $D' = CC$. For I' we proceed as follows.

(i) $\forall\langle c_1, ..., c_n\rangle \in CC^n$, $w \in W$, $P \in \mathcal{P}$:

$$\langle c_1, ..., c_n\rangle \in I'(P)(w) \quad \text{iff} \quad \langle c_1(w), ..., c_n(w)\rangle \in I(P)(w)$$

(ii) $\forall c \in CC$, $w \in W$, $a \in \mathcal{C}$:

$$I'(a)(w) = c \quad \text{iff} \quad I(a)(w) = c(w)$$

In our construction, we take the elements of the conceptual cover in the old model to be the individuals in the new model, and we stipulate that they do, in all w, what their instantiations in w do in the old model. Clause (i) says that a sequence of individuals is in the denotation of a relation P in w in the new model iff the sequence of their instantiations in w is in P in w in the old model. In order for clause (ii) to be well-defined, it is essential that CC is a conceptual cover, rather than an arbitrary set of concepts. In M', an individual constant a will denote in w the unique c in CC such that $I(a)(w) = c(w)$. That there is such a unique c is guaranteed by the uniqueness condition on conceptual covers. We have to prove that this construction works. I will use g, g' for assignments within M and h, h' for assignments within M'. Note that for all assignments g within M: $g(n) = CC$ for all CC-indices n, since CC is the unique cover available in M. I will say that g corresponds with h iff $g = h \cup \{\langle n, CC\rangle \mid n \in N\}$. This means that the two assignments assign the same value to all individual variables x_n for all n, and g assigns the cover CC to all CC-indices n.[50]

2.4.6. Theorem. Let g and h be any corresponding assignments. Let w be any world in W and ϕ any wff in $\mathcal{L}_{CC}$. Then

$$M, w, g \models_{CC} \phi \text{ iff } M', w, h \models_{MPL} \phi$$

Now it is clear that if a classical CC-model M and an ordinary MPL-model M' correspond in the way described, then the theorem entails that any wff in $\mathcal{L}_{CC}$ is CC-valid in M iff it is MPL-valid in M'. Thus, given a classical CC-model, we can define an equivalent MPL-model, but also given an MPL-model, we can define an equivalent classical CC-model $\langle W, R, D, \{CC\}, I\rangle$ by taking CC to be the rigid cover. This suffices to prove proposition 2.4.5.

A corollary of proposition 2.4.5 is that CC-validity is weaker than MPL-validity. $\models_{CC} \phi$ obviously implies $\models_{CCC} \phi$ which by proposition 2.4.5 implies $\models_{MPL} \phi$.

[50] For the complete proof of theorem 2.4.6 see Appendix A.2.

2.4.7. COROLLARY. If $\models_{CC} \phi$, then $\models_{MPL} \phi$.

A further consequence of proposition 2.4.5 is that we can define interesting fragments of $\mathcal{L}_{CC}$ which behave classically, that is, wffs of these fragments are valid iff they are valid in MPL.

Let $\mathcal{L}^n_{CC}$ be a restriction of $\mathcal{L}_{CC}$ containing only variables indexed by n. We can prove the following proposition:

2.4.8. PROPOSITION. $\forall n : \forall \phi \in \mathcal{L}^n_{CC} :\ \models_{CC} \phi \quad \text{iff} \quad \models_{MPL} \phi$

proof: Suppose $\not\models_{CC} \phi$ for $\phi \in \mathcal{L}^n_{CC}$. This means for some CC-model $M = \langle W, R, D, C, I \rangle$ and some w,g: $M, w \not\models_g \phi$. Let $M' = \langle W, R, D, \{g(n)\}, I \rangle$. Since ϕ can only contain variables indexed by n, $M', w \not\models_g \phi$. M' is obviously a classical model. This means $\not\models_{CCC} \phi$ which by proposition 2.4.5 implies $\not\models_{MPL} \phi$. Corollary 2.4.7 delivers the second half of proposition 2.4.8. □

Let $\mathcal{L}_{PL}$ be the non-modal fragment of $\mathcal{L}_{CC}$. We can prove the following proposition:

2.4.9. PROPOSITION. $\forall \phi \in \mathcal{L}_{PL} :\ \models_{CC} \phi \quad \text{iff} \quad \models_{MPL} \phi$

proof: Suppose $\not\models_{CC} \phi$. This means for some CC-model $M = \langle W, R, D, C, I \rangle$ and some w,g: $M, w \not\models_g \phi$. Let $M' = \langle W', R', D, C', I' \rangle$, be a sub-model of M such that $W' = \{w\}$. Since ϕ is non-modal $M', w \not\models_g \phi$. Since $|W'| = 1$, $|C'| = 1$,[51] i.e. M' is a classical model. This means $\not\models_{CCC} \phi$ which by proposition 2.4.5 implies $\not\models_{MPL} \phi$. Again corollary 2.4.7 delivers the other direction of the proof. □

As a consequence of proposition 2.4.9, our CC semantics validates the principles of existential generalization and substitutivity of identicals for non-modal wffs, since they are validated in MPL:

SI1 $\models_{CC} t = t' \rightarrow (\phi[t] \rightarrow \phi[t'])$ (if ϕ is non-modal)

EG1 $\models_{CC} \phi[t] \rightarrow \exists x_n \phi[x_n]$ (if ϕ is non-modal)

Note that the validity of **EG1** crucially relies on the existence condition on conceptual covers, which guarantees that whatever denotation $d = [t]_{M,g,w}$, t is assigned to in w, there is a concept c in the operative cover such that $c(w) = d = [t]_{M,g,w}$.

Substitutivity of identicals and existential generalization cease to hold as soon as we introduce belief operators. By corollary 2.4.7, **SI** and **EG** are invalidated in CC, being invalid in MPL:

¬SI $\not\models_{CC} t = t' \rightarrow (\phi[t] \rightarrow \phi[t'])$

¬EG $\not\models_{CC} \phi[t] \rightarrow \exists x_n \phi[x_n]$

[51] See chapter 4, corollary 4.1.4.

The failures of **SI** and **EG** are welcome because they allow us to solve the *de dicto* substitutivity puzzles (see example (42)) and the shortest spy problems:

(119) $t = t' \not\Rightarrow (\Box Pt \rightarrow \Box Pt')$

(120) $\Box Pt \not\Rightarrow \exists x_n \Box P x_n$

It is easy to show that not only **SI** and **EG** can fail, but also **SIv** and **EGv** are invalidated in the present semantics.

$\neg$ **SIv** $\not\models_{CC} x_n = y_m \rightarrow (\phi[x_n] \rightarrow \phi[y_m])$

$\neg$ **EGv** $\not\models_{CC} \phi[y_m] \rightarrow \exists x_n \phi[x_n]$

From the failure of **SIv**, it follows that also **LIv** is not valid in CC:

(121) $x_n = y_m \not\Rightarrow \Box\, x_n = y_m$

And this allows us to model cases of mistaken identity and to solve the double vision problems.

From the failure of **EGv**, it follows that also the principle of renaming **PR** is not generally valid in CC:

(122) $\exists x_n \Box P(x_n) \not\Leftrightarrow \exists y_m \Box P(y_m)$

And this allows us to deal with the case of Odette's lover and other cases of context sensitivity.

Note finally that substitutivity of identicals and existential generalization are allowed when applied to variables with a uniform index. It is easy to see that the present semantics validates the following schemes:

SIn $\models_{CC} x_n = y_n \rightarrow (\phi[x_n] \rightarrow \phi[y_n])$

EGn $\models_{CC} \phi[y_n] \rightarrow \exists x_n \phi[x_n]$

The validity of **SIn** crucially relies on the uniqueness condition on conceptual covers. From **SIn**, but also as a consequence of proposition 2.4.8, we can derive **LIn**, which guarantees that the elements in our domains of quantification behave more like individuals than representations:

LIn $\models_{CC} x_n = y_n \rightarrow \Box x_n = y_n$

In this section we have seen that modal predicate logic under conceptual covers is essentially richer than standard MPL because we can shift from one cover to another. If we stick to one cover, not only CC and MPL define the same notion of validity (proposition 2.4.5), but also, and maybe more significantly, the same notion of truth (theorem 2.4.6). We have already seen the intuitive consequences of this result. On the one hand, in ordinary situations in which the method of identification is kept constant, CC behaves exactly as MPL and inherits its desirable properties (for instance in relation to the shortest spy problem). On the other hand, the system is flexible enough to account for extraordinary situations as well, such as double vision situations as well as those like Odette's lover which are situations in which multiple covers are operative. So far we have studied the issue of belief attribution from a model theoretic perspective. Let us now turn to the proof theoretic perspective from which modal logic originated. In the next section, I present an axiom system which provides a sound and complete characterization of the set of wffs valid in all CC models.

2.4.5 Axiomatization

In this section, I present the axiom system CC. In Appendix A.2, I prove that the system CC is sound and complete with respect to the class of all CC-models.

The system CC consists of the following set of axiom schemata:[52]

Basic propositional modal system

PC All propositional tautologies.

K $\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$

Quantifiers Recall that $\phi[t]$ and $\phi[t']$ differ only in that the former contains the term t in one or more places where the latter contains t'.

EGa $\phi[t] \to \exists x_n \phi[x_n]$ (if ϕ is atomic)

EGn $\phi[y_n] \to \exists x_n \phi[x_n]$

BFn $\forall x_n \Box\phi \to \Box\forall x_n \phi$

Identity

ID $t = t$

SIa $t = t' \to (\phi[t] \to \phi[t'])$ (if ϕ is atomic)

[52] This axiomatization is based on the axiom system of modal predicate logic with identity in Hughes and Cresswell (1996). See in particular chapters 13, 14 and 17.

SIn $x_n = y_n \to (\phi[x_n] \to \phi[y_n])$

LNIn $x_n \neq y_n \to \Box\, x_n \neq y_n$

Let AX_{CC} be the set of axioms of CC. The set of CC-theorems T_{CC} is the smallest set such that:

AX $AX_{CC} \subseteq T_{CC}$

MP If ϕ and $\phi \to \psi \in T_{CC}$, then $\psi \in T_{CC}$

∃I If $\phi \to \psi \in T_{CC}$ and x^n not free in ψ, then $(\exists x_n \phi) \to \psi \in T_{CC}$

N If $\phi \in T_{CC}$, then $\Box\phi \in T_{CC}$

I will use the standard notation and write $\vdash_{CC} \phi$ for $\phi \in T_{CC}$.

The axioms **EGa** and **SIa** govern existential generalization and substitutivity of identicals for arbitrary singular terms in atomic formulae. **EGn** and **SIn** cover the case for simple variables for general formulae. Note that **EGa** expresses the existence condition on conceptual cover and **SIn** the uniqueness condition.

In atomic contexts, existential generalization is applicable to any term (**EGa**), and any two co-referential terms are interchangeable *salva veritate* (**SIa**). This can be generalized to any non-modal context. In the CC system, we can deduce **EG1** and **SI1**:[53]

EG1 $\vdash_{CC} \phi[t] \to \exists x_n \phi[x_n]$ (if ϕ is non-modal)

SI1 $\vdash_{CC} t_1 = t_2 \to (\phi[t_1] \to \phi[t_2])$ (if ϕ is non-modal)

On the other hand, any n-indexed variable occurring in any arbitrary context is suitable for n-existential generalization (**EGn**), and any two co-referring variables indexed in a uniform way can be substituted *salva veritate* in any context (**SIn**).

There is another pair of related theorems derivable in CC, which govern existential generalization and substitutivity of identicals for formulae with one layer of modal operators:[54]

[53]**SI1** may be deduced from **SIa** by induction on the construction of $\phi[t_1]$ and $\phi[t_2]$ (the proof is standard). From **SI1** we can derive **EG1** as follows (for ϕ non-modal):

(1)	$\vdash_{CC} t = x_n \to (\phi[t] \to \phi[x_n])$	**SI1**
(2)	$\vdash_{CC} t = x_n \to (\phi[t] \to \exists x_n \phi[x_n])$	(1)×**EGn**×**PC**
(3)	$\vdash_{CC} \exists x_n(t = x_n) \to (\phi[t] \to \exists x_n \phi[x_n])$	(2)×**∃I**
(4)	$\vdash_{CC} \exists x_n(t = x_n)$	**ID**×**EGa**×**MP**
(5)	$\vdash_{CC} \phi[t] \to \exists x_n \phi[x_n]$	(4)×(3)×**MP**

[54]$\mathbf{SI}_\Box$ may be deduced from **SI1**, **N** and **K**. From $\mathbf{SI}_\Box$, we may derive $\mathbf{EG}_{\Box n}$ as follows (for ϕ non-modal):

EG$_{\Box n}$ $\vdash_{CC} \exists x_n \Box t = x_n \rightarrow (\Box\phi[t] \rightarrow \exists x_n \Box\phi[x_n])$ (if ϕ is non-modal)

SI$_{\Box}$ $\vdash_{CC} \Box t_1 = t_2 \rightarrow (\Box\phi[t_1] \rightarrow \Box\phi[t_2])$ (if ϕ is non-modal)

If we add to our axiomatic base the principles **D**, **4** and **E**, these two theorems can be generalized to any ϕ as we expect to be the case for a logic of belief.

Finally note that **BFn** and **LNIn** have the property that they are derivable for some other choices of the basic propositional modal system, e.g. B or S5. I will not consider those systems though, because **B** is not a plausible principle for a logic of belief, so we have to take **BFn** and **LNIn** as axioms. **LIn** is instead derivable in CC, as well as the n-versions of the converse of the Barcan Formula and the principle of importation. The proofs are standard.

LIn $\vdash_{CC} x_n = y_n \rightarrow \Box\, x_n = y_n$

CBFn $\vdash_{CC} \Box\forall x_n\phi \rightarrow \forall x_n\Box\phi$

IMn $\vdash_{CC} \exists x_n\Box\phi \rightarrow \Box\exists x_n\phi$

The converse of **IMn**, instead, is *not* provable and it is not valid indeed.

¬**EXn** $\nvdash_{CC} \exists x_n\Box\phi \rightarrow \Box\exists x_n\phi$

In Appendix A.2, we prove that the system CC is sound and complete with respect to the set of all CC-models.

2.4.10. Theorem. [Soundness] If $\vdash_{CC} \phi$, then $\models_{CC} \phi$.

2.4.11. Theorem. [Completeness] If $\models_{CC} \phi$, then $\vdash_{CC} \phi$.

By standard techniques we can show that CC+**D**+**4**+**E** is sound and complete with respect to all serial, transitive and euclidean CC-models.

2.5 Synopsis

The following diagram summarizes the content of this chapter. On the topmost horizontal row, the four systems are displayed that we have encountered in the previous sections. On the second column from the left, the principles we have discussed are listed; on the leftmost column, the problems are reported, which are caused by the validity of these principles. The $\models_{(*)}$ or $\not\models_{(*)}$ indicate that the relevant systems do or do not validate the corresponding principles and that this is problematic.

---

(1) $\vdash_{CC} \Box t = x_n \rightarrow (\Box\phi[t] \rightarrow \Box\phi[x_n])$ **SI**$_{\Box}$

(2) $\vdash_{CC} \Box t = x_n \rightarrow (\Box\phi[t] \rightarrow \exists x_n\Box\phi[x_n])$ (1) ×**EGn**× **PC**

(3) $\vdash_{CC} \exists x_n\Box t = x_n \rightarrow (\Box\phi[t] \rightarrow \exists x_n\Box\phi[x_n])$ (2) ×∃**I**

		MPL	CIA	CIB	CC
de dicto substitutivity puzzles	**SI**	$\not\models$	$\not\models$	$\not\models$	$\not\models$
	SI1	$\models$	$\models$	$\models$	$\models$
(*) double vision problems **LIv**	**SIv**	$\models_{(*)}$	$\not\models$	$\not\models$	$\not\models$
	SIn				$\models$
(*) shortest spy problems **(T)EX**	**EG**	$\not\models$	$\models_{(*)}$	$\not\models$	$\not\models$
	EG1	$\models$	$\models$	(?) $\not\models_{(*)}$	$\models$
(*) Odette's lover and other problems **PR**	**EGv**	$\models_{(*)}$	$\models_{(*)}$	$\models_{(*)}$	$\not\models$
	EGn				$\models$

We started by discussing the principles of substitutivity of identical **SI** and existential generalization **EG**. **SI** fails in all considered systems, once they interpret individual constants as non-rigid designators. The failure of **SI** allows us to avoid the *de dicto* substitutivity puzzles. The difficulty of MPL was that it failed to account for the dependence of belief on the ways of specifying objects and, therefore, it ran in the double vision problems (by verifying **SIv**, MPL verifies **LIv**). The CIA solution to these problems consisted in letting variables range over all individual concepts rather than all objects (**SIv** is falsified in CIA). However such a strategy led directly to the shortest spy problems (principles **EG** and **(T)EX** are validated in CIA). CIB solved both problems by letting variables range over suitable subsets of the set of all individual concepts (**SIv** and **EG** are not CIB-valid). But, since the information about the suitable concepts was determined by the model, rather than by a contextual parameter, the system could not avoid the problem of Odette's lover and in general could not account for the context sensitivity of *de re* constructions (**EGv** and so **PR** are validated in CIB). Furthermore, without further refinement, CIB does invalidate **EG1**, which is highly counter-intuitive (see discussion around **EG1** in section 2.3.3).

The CC analysis solves these problems by staying as close as possible to MPL. In CIA and CIB, variables range over sets not governed by the principle of substitutivity of identicals (**EGv** holds, whereas **SIv** fails), so typically over sets of representations. On the other hand, in MPL and CC, the 'objects' over which we quantify, cannot split once we move from one world to the other (**EGv** and **SIv** hold in MPL, and **EGn** and **SIn** hold in CC), and therefore behave like individuals, rather than representations of individuals. But while in MPL, the validity of **SIv** and **EGv** led to the double vision and the problem of Odette's lover respectively, in CC, only the weaker **SIn** and **EGn** are validated. **SIv** and **EGv** can fail and, therefore, cases of mistaken identity and of context sensitivity can be accounted for.

2.6 Conclusion

Many authors have recognized the availability of different methods of cross-identification, and argued that in different contexts different methods can be used. The present analysis was an attempt to give a precise formalization of this insight and to discuss its impact on the interpretation of *de re* belief attributions. By taking variables to range over elements of contextually selected conceptual covers, we account for the ordinary sense of belief, according to which belief attributions depend on ways of specifying objects, while avoiding the counterintuitive results which arise when we quantify over ways of specifying individuals rather than over the individuals themselves.

Chapter 3

Dynamics

In dynamic semantics, different styles of quantification have been proposed that involve two different ways of interpreting free and quantified variables:

(i) Variables as denoting single partial objects;

(ii) Variables as ranging over a number of alternative total objects.

In the first part of the present chapter, I show that the first view leads to problems of underspecification and the second to problems of overspecification. In the second part, I propose a new style of dynamic quantification in which variables are interpreted in a way that avoids these problems:

(iii) Variables as ranging over a number of alternative definite objects (concepts).

I will then show that specific problems which arise when we quantify over concepts rather than objects, are avoided by relativizing quantification to ways of conceptualizing the domain.

3.1 Dynamic Semantics

In dynamic semantics,[1] the formal meaning of a natural language expression is identified with its potential to change an information state. An information state is generally characterized as a set of possibilities, consisting of the alternatives which are compatible with the information of the relevant agents. The nature of these possibilities depends on what particular aspect of the information change potential of a sentence one studies, and this is relative to the kind of phenomena

[1]Dynamic semantics originates from Kamp (1981), Heim (1983a), Groenendijk and Stokhof (1991) and Kamp and Reyle (1993). See Dekker (1993) and van Benthem et al. (1997) for excellent overviews.

one is willing to account for. In the present chapter, I study the interaction between anaphora and notions like epistemic modalities or presupposition, and, therefore, the type of information at issue concerns the state of the world, and what are possible antecedents for anaphoric pronouns.

Anaphora constitutes the traditional area of application of dynamic semantics (see Kamp (1981), Heim (1983a), Groenendijk and Stokhof (1991), Chierchia (1992), Dekker (1993), and others). Consider the following classical examples of inter-sentential and donkey-anaphora:

(123) I met a woman last night. She was feeding pigeons in the park.

(124) If a farmer has a donkey, he is rich.

These examples constitute a problem for a classical montagovian semantics since their arguably compositional logical representations in (125) and (126) do not reflect the intuitive meanings of the sentences, if interpreted in a standard fashion, since the variable x in the second conjunct in (125) and in the consequent in (126) occurs outside the syntactic scope of the existential quantifier $\exists x$ and hence it is not bound by it.

(125) $\exists x\phi_1(x) \wedge \phi_2(x)$

(126) $\exists x\psi_1(x) \rightarrow \psi_2(x)$

Dynamic semantics manages to solve these difficulties by encoding, as part of the meaning of indefinite NPs, their potential to introduce new items which can serve as antecedents for subsequent anaphora. Information about a potential antecedent is characterized as information about the possible values of a variable. The main feature of the dynamic existential quantifier is that it can bind variables outside its syntactic scope. The following are valid dynamic equivalences:

(127) $\exists x\phi(x) \wedge \psi(x) \equiv \exists x(\phi(x) \wedge \psi(x))$

(128) $\exists x\phi(x) \rightarrow \psi(x) \equiv \forall x(\phi(x) \rightarrow \psi(x))$

In what follows, I will use the traditional terminology and call *quantified* variables, variables occurring in the syntactic scope of a quantifier and *free* variables, variables occurring outside the syntactic scope of a quantifier. Crucially, such free occurrences may still be *dynamically bound* by a quantifier.

Epistemic modalities (see Veltman (1997)) and presupposition (Heim (1983b), Beaver (1995), van der Sandt (1992), Chierchia (1995) and others) constitute another traditional application area for a dynamic approach. Consider the following well-known examples:

(129) a. Someone is knocking at the door. It might be Mary. ... It is John.

b. Someone is knocking at the door. It is John. ... (?) It might be Mary.

(130) a. Bill likes Mary and John likes Mary too.

b. (?) John likes Mary too and Bill likes Mary.

Dynamic systems have been proposed, which view meanings as potentials to change factual information, and are thus able to capture the contrast between the (a) and (b) texts in (129) (see Veltman (1997)) and the projection of presupposition in conjunctions illustrated in (130) (e.g. Heim (1983b)). A characteristic, technical feature of a dynamic system, which is illustrated by these cases, is the fact that dynamic conjunction is not commutative:

(131) $\phi \wedge \psi \not\equiv \psi \wedge \phi$

Like we said, information about potential antecedents for future anaphora can be encoded by means of assignment functions (see Heim (1983a), Groenendijk and Stokhof (1991), etc.). Factual information can be represented by means of possible worlds (see Veltman (1997), Heim (1983b), etc.). A dynamic systems which sets out to investigate the interaction between anaphora, on the one hand, and epistemic modality and presupposition on the other, can characterize information states as sets of world-assignment pairs (see Heim (1982), Dekker (1993), etc.). The issue of the proper combination of these two kinds of information constitutes the main theme of the present chapter.

3.2 Quantification in Dynamic Semantics

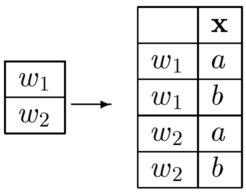
In a dynamic system, sentences describe transitions across a space of information states. As we said in the previous section, information states are defined as sets of possibilities (here world-assignment pairs) and meanings are state transitions, that is, functions or relations over the space of information states. An update with a sentence may reduce the size of a state or may yield richer states. Atoms or negations narrow down the alternatives under consideration by eliminating the world-assignment pairs that do not satisfy the information contents of these sentences. Existentially quantified sentences add structure to the state by setting up new items as potential topics for further discourse: $\exists x\phi$ adds x and selects a number of possible values for it; the fact that in the output state(s) x is defined means that recurrences of x in later sentences can have the effect of anaphoric reference.

Information about the values of variables is generally modeled in one of the following two ways:

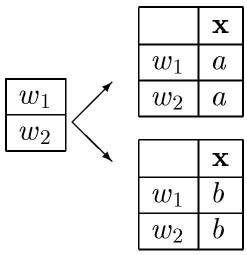
1. Variables are interpreted as single *partial* objects.[2] The introduction of new items is defined in terms of *global extensions* that involve adding fresh variables and assigning them as possible values all elements of the universe of discourse. All of the values which variables can take are considered simultaneously.

2. Variables are taken to range over a number of *total* objects. The introduction of a new item is defined in terms of *individual extensions* that lead to states in which the added variable is assigned a single element of the universe as its value. The values which variables can take are considered one by one as disjoint alternatives.

Individual and global extensions can be depicted as follows:[3]



1. Global Extension



2. Individual Extension

Global extensions yield unique output states, whereas individual extensions produce as many different outputs as there are members of the universe. This involves splitting up the initial state into different alternatives: later sentences will be interpreted with respect to each of them in a parallel fashion.

In the literature, three different interpretations have been proposed for the dynamic existential quantifier and they involve one or the other way of interpreting free and quantified variables:[4]

[2]Partial objects are the structured entities that constitute the interpretations of variables in information states. In Dekker (1993), they are defined as functions that assign to each possibility in a state the value of the corresponding variable in that possibility. A partial object is called total if it is a constant function. In the picture above, the partial object corresponding to the interpretation of the variable x is represented by the vertical column below x.

[3]I use these pictures to represent shifts on information states. The tables correspond to information states. On the topmost horizontal row, the variables that are defined in the state are displayed in bold characters. Each other horizontal row represents a world-assignment element of the state. The left column contains the world-coordinate and to its right the values of the assignment functions with each value displayed right below the variable which it gets assigned to. In this picture, the universe is assumed to consist of only two individuals a and b.

[4]Theoretically a fourth possibility is conceivable, according to which quantified variables are interpreted as ranging over total objects and free variables receive a partial interpretation. As will be clear from the following discussion, this possibility makes no intuitive sense, and, likewise, it has never been proposed in the literature.

Random Assignment (RA) is the standard interpretation. It is defined in terms of global extension; in RA fresh variables are assigned all individuals from the universe of discourse as possible values. In this way, quantified and free variables are interpreted uniformly as single indefinite partial objects, where further updates tend to make these objects more definite and less partial. See Heim (1982), Heim (1983b), Dekker (1993).

Slicing (SL) is defined in terms of individual extension; it involves splitting up the update procedure, so that the values that a variable can have are considered one by one, as disjunct alternatives, and not all at once. In this way, quantified and free variables are interpreted uniformly as ranging over a number of alternative total objects, where further updates tend to eliminate certain alternatives. See van Eijck and Cepparello (1994).

Moderate Slicing (MS) follows the slicing procedure as long as we are inside the syntactic scope of a quantifier, but lumps the remaining alternatives together once we are outside its scope. In this way, quantified variables range over a number of alternative total objects, whereas free variables are interpreted as single partial objects. See Beaver (1994), Dekker (1994), Groenendijk et al. (1996).[5]

These different styles of quantification lead to different results only in connection with notions that are sensitive to global properties of information states, i.e., notions that take a state as a whole and not point-wise with respect to the possibilities in it. This is not surprising: if we take states holistically it is obvious that it matters which possibilities are lumped together to form a state and which are kept separate. Examples of holistic notions are epistemic modals,[6] presupposition,[7] and the notion of support.[8]

Although the analysis of combinations of quantifiers and holistic notions motivated the use of (moderate) slicing instead of random assignment, I will argue that precisely in such contexts critical problems emerge for all three styles of quantification. Before turning to the illustration of these problems, let me introduce the relational dynamic semantics that supplies the general framework for the comparison of the three approaches.

Formal Framework

The core of the semantic framework that I take as a starting point is a relational version of the update semantics MDPL presented in Dekker (1993), with the ad-

[5]After the article has been written, upon which this chapter is based, Zuidema (1999) has proposed a system in which variables are introduced by individual extensions, but their possible values can be lumped together at later stages of the interpretation by means of a *collapse* operator. A proper discussion of this very interesting approach must be left to another occasion.

[6]See Dekker (1993), van Eijck and Cepparello (1994), Groenendijk et al. (1996), Veltman (1997).

[7]See Heim (1983b), Beaver (1994), Beaver (1995).

[8]See Groenendijk et al. (1996) and Dekker (1997).

dition of the presupposition operator introduced in Beaver (1995). The *language* $\mathcal{L}$ is a standard predicate logical language with the addition of two sentential operators, the epistemic modal operator $\Diamond$ and the presupposition operator ∂. Given $\mathcal{L}$, a *model* M for $\mathcal{L}$ is a pair $\langle W, D\rangle$ where W, the set of possible worlds, is a non-empty set of interpretation functions for the non-logical constants in $\mathcal{L}$, and D, the domain of discourse, is a non-empty set of individuals. Information states are sets of possibilities. They are defined as in Heim (1982) and Dekker (1993) as sets of world-assignment pairs in which all the assignment functions have the same domain.

3.2.1. Definition. [Information States] Let $M = \langle D, W\rangle$ be a model for $\mathcal{L}$. Let $\mathcal{V}$ be the set of individual variables in $\mathcal{L}$. The set Σ_M of *information states* based on M is defined as:

$$\Sigma_M = \bigcup_{X \subseteq \mathcal{V}} \mathcal{P}(W \times D^X)$$

I will use I^X to denote the set of possibilities $W \times D^X$ for some $X \subseteq \mathcal{V}$, and if $i = \langle w, a\rangle$ is a possibility, I will write w_i for w and a_i for a.

A possibility in an information state contains enough information for the interpretation of the basic expressions in $\mathcal{L}$.

3.2.2. Definition. Let α be a basic expression in $\mathcal{L}$ and i a possibility in I^X for some $X \subseteq \mathcal{V}$. *The denotation* of α in i is defined as:

(i) if α is a non-logical constant, then $i(\alpha) = w_i(\alpha)$;

(ii) if α is a variable in X, then $i(\alpha) = a_i(\alpha)$, undefined otherwise.

Meanings are relations over Σ_M. Before stating the semantic clauses, we need to define the auxiliary notion of survival.

Survival is a relation between a possibility and an information state and, indirectly, between two information states (cf. Dekker (1993)).

3.2.3. Definition. [Survival] Let $\sigma, \sigma' \in \Sigma_M$ & $i \in I^X$ for some $X \subseteq \mathcal{V}$. Then

(i) $i \prec \sigma \quad \text{iff} \quad \exists j \in \sigma : w_i = w_j \;\&\; a_i \subseteq a_j$;

(ii) $\sigma \prec \sigma' \quad \text{iff} \quad \forall i \in \sigma : i \prec \sigma'$.

A world-assignment pair i survives in a state σ iff σ contains a possibility j such that j is the same as i except for the possible introduction of new variables. A state σ survives in a state σ' iff all possibilities in σ survive in σ'.

We can now turn to the simultaneous definition of the main semantic clauses and of the notion of update and support in our system. (In this first definition, we skip the interpretation of $\exists\phi$ which will be discussed shortly.)

3.2.4. Definition. [Support] Let $\sigma \in \Sigma_M$ and ϕ in $\mathcal{L}$. Then

$$\sigma \approx\!\!\!\mid\; \phi \quad \text{iff} \quad \exists\sigma' :\ \sigma[\phi]\sigma' \ \&\ \sigma \prec \sigma'$$

A state σ supports a sentence ϕ iff all possibilities in σ survive simultaneously in at least one of the states resulting from updating σ with ϕ, where updates are defined as follows:

3.2.5. Definition. [The Core of the Semantics]

$$\begin{array}{rll}
\sigma[Rt_1, ..., t_n]\sigma' & \text{iff} & \sigma' = \{i \in \sigma \mid \langle i(t_1), ..., i(t_n)\rangle \in i(R)\}; \\
\sigma[\neg\phi]\sigma' & \text{iff} & \sigma' = \{i \in \sigma \mid \neg\exists\sigma'' : \sigma[\phi]\sigma'' \ \&\ i \prec \sigma''\}; \\
\sigma[\phi \wedge \psi]\sigma' & \text{iff} & \exists\sigma'' :\ \sigma[\phi]\sigma''[\psi]\sigma'; \\
\sigma[\Diamond\phi]\sigma' & \text{iff} & \sigma' = \{i \in \sigma \mid \exists\sigma'' \neq \emptyset : \sigma[\phi]\sigma''\}; \\
\sigma[\partial\phi]\sigma' & \text{iff} & \sigma \approx\!\!\!\mid\; \phi \ \&\ \sigma[\phi]\sigma'.
\end{array}$$

Updating a state σ with an atomic formula preserves those possibilities in σ which satisfy the formula in a classical sense. The negation of ϕ eliminates those i in σ which can survive after updating σ with ϕ. Conjunction is relational composition.

Modal sentences $\Diamond\phi$ are interpreted in Veltman's style, as consistency tests.[9] Updating with $\Diamond\phi$ involves checking whether ϕ is consistent with the information encoded in the input state σ. If the test succeeds, i.e., if at least one world-assignment pair in σ survives an update with ϕ, then the resulting state is σ itself, so nothing happens; if the test fails, the output state is the empty set, i.e. the absurd state (cf. Veltman (1997)).

∂ is Beaver's presupposition operator.[10] $\partial\phi$ should be read as 'it is presupposed that ϕ' and is interpreted as an update that is defined on a state σ only if ϕ is already supported in σ. Notice that presuppositions are not simple tests — the output state may vary from the input state in that it can contain new discourse items (cf. Beaver (1995)).

Consistency tests, presupposition and support are holistic notions because they relate to properties of the whole state, not of its individual elements.

Three different systems of interpretation can be developed from this core semantics depending on which of the three above-mentioned forms of dynamic existential quantification we adopt. To define them we need to introduce the auxiliary

[9] Many aspects of the meaning of epistemic modals in natural language are not captured by this analysis (see Roberts (1996a) for discussion). However, the one aspect that is addressed, namely that upon hearing *It might be that* ϕ one checks whether one's information is consistent with the information contained in ϕ, is significant and has a non-distributive nature.

[10] The empirical adequacy of Beaver's presupposition operator has been discussed (see van der Sandt (1992) and Geurts (1996) for discussion). However, the aspect that is captured by this definition, namely that an utterance of *John regrets that* ϕ is infelicitous unless the background state already supports the information that ϕ, is significant and has a non-distributive nature.

notions of an assignment operation, a global extension and an individual extension.

Assignment operations extend possibilities by adding fresh[11] variables and assigning to them as values individuals from the domain.

3.2.6. DEFINITION. [Assignment Operations] Let $i \in I^X$ for some $X \subseteq \mathcal{V}$, $x \notin X$ and $d \in D$. Then

$$i[x/d] = \langle w_i, a_i \cup \{\langle x, d\rangle\}\rangle$$

We can now define both global and individual extensions.

3.2.7. DEFINITION. [Extensions] Let $\sigma \subseteq I^X$, $x \notin X$ and $d \in D$. Then

(i) $\sigma[x] = \{i[x/d] \mid d \in D \;\&\; i \in \sigma\}$ (*global*);

(ii) $\sigma[x/d] = \{i[x/d] \mid i \in \sigma\}$ (*individual*).

Global extensions add fresh variables and randomly assign *all* elements of the universe of discourse to them. Individual extensions enlarge the domain of the state by assigning *single* elements of D to fresh variables. In terms of the notion of an extension, we can finally define Random Assignment, Slicing and Moderate Slicing.

3.2.8. DEFINITION. [Three Styles of Quantification]

$$\begin{array}{rll} \sigma[\exists x\phi]_{RA}\sigma' & \text{iff} & \sigma[x][\phi]\sigma'; \\ \sigma[\exists x\phi]_{SL}\sigma' & \text{iff} & \sigma[x/d][\phi]\sigma' \text{ for some } d \in D; \\ \sigma[\exists x\phi]_{MS}\sigma' & \text{iff} & \sigma' = \cup_{d\in D}\{\sigma'' \mid \sigma[x/d][\phi]\sigma''\}. \end{array}$$

Universal quantification is defined in the standard way in terms of negation and existential quantification.

Since all of the operations defined are functional with the only exception of SL, if either MS or RA are assumed as the interpretation of the existential quantifier, then the whole semantics can be stated in terms of (partial) functions.

We can now turn to the illustrations of the problems.

[11] As in Heim (1982) and Dekker (1993), variables cannot be reset because resetting variables would involve losing information about their previous values. This 'downdate' effect would be problematic for the notions of negation and support, which crucially rely on the fact that no operations are considered that cause loss of information. There are other means, though, to avoid the 'downdate' problem, which allow reuse of variables, see for instance Groenendijk et al. (1996), Vermeulen (1996) and Dekker (1996). Notice that once we assume the style of quantification I eventually propose, we can reformulate the semantics in such a way that downdates are no longer problematic (cf. section 3.6). Finally observe that the novelty condition is a source of partiality. In addition to (i) presuppositions and (ii) formulas containing free variables, (iii) quantified sentences are partial updates as well. In the system I will present in section 3.6, only presupposition can cause undefinedness. Since partiality introduced by presupposition is not directly relevant to the issues discussed in the chapter, and partiality (ii) and (iii) do not occur in the final version of my system, I will pass over the issue of undefinedness in what follows.

3.3 Underspecification and Overspecification

In this section I will argue that of the two ways of interpreting variables that play a role in the three styles of dynamic quantification, the one that treats variables as single partial objects is too weak and leads to problems of *underspecification*. The other, which views them as place-holders for a number of alternative total objects, is too strong and leads to problems of *overspecification*.

Underspecification 1

If *quantified* variables are interpreted as *partial* objects, difficulties arise in connection with phenomena that involve quantification into the scope of holistic operators. Consider the following examples.

the suspect Treating variables in the syntactic scope of a quantifier as single underspecified objects has the unfortunate consequence that for all states σ the following holds (cf. Dekker (1993)):

(132) $\sigma \models\!\!\approx \exists x \Diamond \phi \quad \Rightarrow \sigma \models\!\!\approx \forall x \Diamond \phi$

So if we assume RA, sentences like the following two will contradict each other:

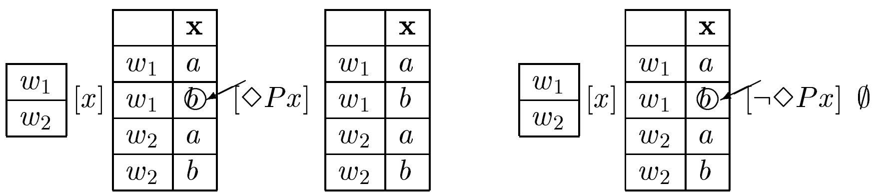
(133) a. Someone might be the culprit.

b. $\exists x \Diamond Px$

(134) a. Someone certainly is not the culprit.

b. $\exists x \neg \Diamond Px$

However, intuitively (133) and (134) express compatible pieces of information: you may hold the guilt of someone to be consistent with your information state and at the same time have evidence that someone else is innocent. The problem with RA is that the variable x introduced via global extension denotes exactly the same single underspecified object in both cases, which either verifies the modal sentence $\Diamond Px$, or falsifies it.



$w_1(P) = \{b\}$

If at least one member of the universe has the property P in some world (say individual b in world w_1 as in the picture), (133) is accepted and (134) is rejected. If this is not the case the opposite holds. So (133) and (134) cannot be accepted at the same time. This undesirable result obtains because, to quote from Beaver (1994), in *RA*, quantified variables don't vary enough: the one value that a variable can take cannot be considered separately from the others, because all the possible values are lumped together. This type of underspecification is also the source of the problem discussed in the following example.

the fat man This problem discussed in Heim (1983b) concerns the projection of presuppositions from quantified contexts. Consider (135):

(135) a. A fat man was pushing his bicycle.

b. $\exists x[\text{fat-man}(x) \wedge \partial(\exists y\ \text{bike-of}(x, y)) \wedge\ \text{pushing}(x, y)]$

Intuitively, (135) projects the presupposition that the intended fat man had a bicycle.[12] However, Heim (1983b), which assigns a random interpretation to variables, predicts the universal presupposition *Every fat man has a bicycle* for sentence (135), which is too strong intuitively.[13] The ∂ clause is interpreted with respect to the state resulting from adding x and updating with fat-man(x). If x is introduced by Random Assignment, this local state may contain several alternative values of x for each surviving world, namely all fat men in that world (a and b in the picture below). If any of these values is not a bike owner, then the ∂ clause turns out undefined. That is, in each world all fat men (all possible values of x) must own a bike, otherwise the sentence is not accepted.

$$\begin{array}{|c|}\hline w_1 \\ \hline w_2 \\ \hline \end{array}\ [x] \circ [\text{fat-man}(x)]\ \begin{array}{|c|c|}\hline & \mathbf{x} \\ \hline w_1 & a \\ \hline w_1 & b \\ \hline w_2 & a \\ \hline w_2 & b \\ \hline \end{array}\ [\partial(\exists y\ \text{bike-of}(x, y))]$$

[12] In Karttunen and Peters (1976), (135) is predicted to have the existential presupposition *Some fat man had a bicycle*. This prediction, as the authors admit, is clearly too weak, because intuitively, what should be projected in this case is the presupposition that the *same* fat man that verifies (135) had a bicycle, and *not* some other fat man. The problem arises because in K&P's system there is no obvious way to define scope and binding relations between the presupposition and the assertion, since these two components are represented by two mutually independent propositions. Note, however, that in dynamic semantics or DRT in which variables in one proposition can be bound by quantifiers in another proposition this problem does not occur. See Dekker (1998).

[13] In the same paper, Heim suggests remedying this inadequacy by stipulating the ready availability of an *ad hoc* accommodation mechanism in the evaluation of indefinite sentences. Standard accommodation mechanisms do not apply, because the relevant presupposition here must be accommodated locally.

Like Dekker's problem, Heim's problem results from the fact that in holistic updates all of the values that variables can take are considered all at once instead of one at a time.

Overspecification 1

If we use slicing, the two problems discussed above do not occur.[14] However, the *total* interpretation of *free* variables that SL involves, leads to the loss of a number of attractive properties guaranteed by MS in connection with phenomena of identification in situations of partial information.

the culprit Consider the following examples discussed by Groenendijk et al. in (1996) that involve dynamically bound variables occurring in the scope of Veltman's epistemic operator:

(136) a. Someone did it. It might be you. It might also not be you.[15]

b. $\exists x Px \wedge \Diamond(x = \text{you}) \wedge \Diamond(x \neq \text{you})$

(137) a. Someone did it. It might be anyone.

b. $\exists x Px \wedge \forall y \Diamond(x = y)$

These are coherent pieces of discourse, but if variables range over alternative total objects, they are both inconsistent. Take for example (137), which expresses an ultimate form of ignorance about the culprit's identity. If variables are place-holders for individuals, updating with (137) always yields the absurd state since it is impossible for one individual to be (possibly) identical to all the others (if $|D| > 1$). In RA and in MS, in which free variables are viewed as partial objects (136) and (137) are instead coherent, as should be the case.

Underspecification 2

The use of moderate slicing avoids the problems noted above, but runs into several others connected with the notions of presupposition, support and coherence. The source of the difficulties here is MS's *partial* interpretation of *free* variables.

[14] An alternative solution to underspecification 1 is obtained by defining presupposition (cf. Beaver (1992)) and modality (cf. Beaver (1993)) in a different way. However, by adopting (moderate) slicing (cf. Beaver (1994) and Groenendijk et al. (1996)), we obtain the same results with minor surgery.

[15] In this example, the deictic pronoun *you* is assumed to rigidly refer to the same individual in all epistemic possibilities.

the fat man again Heim's fat man problem arises not only for quantified variables, but for free variables as well. As an illustration, consider the following variation of (135), in which the occurrence of the variable x in the ∂ clause is dynamically bound by the existential quantifier:

(138) a. A fat man was sweating. He was pushing his bicycle.

b. $\exists x[\text{fat-man}(x) \wedge \text{ sweat}(x)] \wedge \partial(\exists y \text{ bike-of}(x, y)) \wedge \text{ pushing}(x, y)$

If we assume a partial interpretation of free variables (RA and MS), then, for the same reasons as above, (138) projects the presupposition that every fat man who was sweating had a bike, which is intuitively too strong.[16]

the wrong suspect Further difficulties for the partial view of free variables arise in connection with the notions of support and coherence. The notion of support (see definition 3.2.4) can be used to characterize when a speaker is licensed to utter a certain proposition. A speaker is licensed to utter ϕ iff her own information state supports ϕ. As a straightforward generalization we may say that a sentence is assertable iff there is a non-absurd state that supports it. In Groenendijk et al. (1996), texts satisfying this condition are called coherent texts; intuitively, such texts express mutually compatible pieces of information. Now, consider the following example (the pronoun in the second sentence should be read as co-referential with the indefinite in the first sentence):

(140) a. Someone might be the culprit. She is not the culprit.

b. $\exists x \Diamond Px \wedge \neg Px$

Intuitively (140) cannot be coherently asserted as a continuous monologue.[17] The first and second sentence express incompatible pieces of information. You cannot hold the guilt of a person to be consistent with your information and at the same time have the information that the same person is innocent. But if we use MS (or RA) and treat free variables as denoting single partial objects, (140) surprisingly comes out coherent, i.e. there are states that support it. Take as input a state σ^* consisting of two possibilities that supports the information that either individual a (in w_2) or individual b (in w_1) is P. In MS (as in RA), which

[16]In contrast with the original fat man case, standard accommodation could be used in this case. However, this does not improve the situation for an advocate of MS who would then have to explain why the free variable case (138) requires accommodation whereas the following bound variable variation does not:

(139) A fat man who was sweating was pushing his bicycle.

Intuitively, the two problems should be solved in a single move.

[17]Of course, you might be licensed to utter such a conjunction if there is a break between the two conjuncts. In the break you might have gained new information.

allows a partial interpretation of free variables, the first conjunct leads to a state with four possibilities in which both a and b are assigned as possible values to x for each world. Updating with the second conjunct keeps only those two possibilities that assign to x the individuals that are not P:

$$\begin{array}{|c|}\hline w_1\\\hline w_2\\\hline\end{array}\ [\exists x \Diamond Px]\ \begin{array}{|c|c|}\hline & \mathbf{x}\\\hline w_1 & a\\\hline w_1 & b\\\hline w_2 & a\\\hline w_2 & b\\\hline\end{array}\ [\neg Px]\ \begin{array}{|c|c|}\hline & \mathbf{x}\\\hline w_1 & a\\\hline w_2 & b\\\hline\end{array}$$

Even though the latter update eliminates possibilities, both possibilities in the initial state survive in the final state. So σ^* supports the sequence and hence the latter is coherent. It is impossible, however, for a state resulting from a successful update with the first sentence in (140) to support the second one. The fact that (140) still comes out coherent shows the 'non-compositionality' of the notion of support: we may have a state that supports a conjunction, whereas the same state updated with the first conjunct does not support the second one. The notion of support predicts that a speaker who is licensed to assert $\phi_1 \wedge \phi_2$ as a whole, is not necessarily licensed to assert ϕ_2 after asserting ϕ_1 and this is counter-intuitive.[18]

To summarize, in both MS and RA, in which free variables are interpreted as single partial objects, texts like (138) are predicted to project too strong universal presuppositions, texts like (140) come out counter-intuitively coherent, and, connected with this, we have a 'non-compositional' notion of support.

Overspecification 2

The *total* interpretation of *quantified* or free variables hides the conceptual presupposition that there exists a unique method of individuation across the boundaries of our epistemic possibilities. In Groenendijk et al. (1996), the total objects in an information state are taken to represent the ordinary individuals the agents are acquainted with and, following the 'russellian' tradition, these are specified as objects of perception. Given our trust in our perceptual capacities, it is quite reasonable to assume that if an individual is standing in front of us, then the *same* individual will be standing in front of us in all our epistemic alternatives. So demonstrative identification, as opposed to descriptive identification, is suggested as the unique correct method of cross-identification, and direct reference, that is, reference to the 'objects themselves', is specified as reference under such

[18] See Dekker (1997) and more recently Dekker (2000b), in which this problem is solved by introducing a new notion of support. All underspecification problems can be solved in RA by adopting different analyses for the three holistic notions. However, if underspecification can be avoided by simply using another style of quantification, then by dropping RA we account for three groups of phenomena by means of a single move.

a perspective. The problem with this characterization is that it fails to account for phenomena of identity and identification in situations of partial or mistaken information, that are precisely the kind of phenomena that quantified epistemic logic should account for.[19]

the man with the hood Suppose a man with a hood is standing in front of you and you haven't the faintest idea who he is. Groenendijk et al. have no obvious way of expressing this uncertainty. The following natural candidate, for instance, comes out inconsistent:

(141) $\neg\exists x\Box(x = \text{this})$

If variables range over total objects, (141) is accepted in a state iff the semantic value of the demonstrative is *not* a total object. The problem is that, if demonstrative identification is taken as the unique identification method, then that same man is standing in front of you with a hood on his head in all your epistemic alternatives, and so, by definition, you have identified him.

We could conclude that demonstrative identification was not the right characterization of identity across epistemic possibilities and that we should look further for a more adequate notion. However, it is easy to see that this would not be the right way to go. Similar problematic cases can be constructed for any other possible characterization of such a notion. Direct reference cannot be characterized as reference to the objects identified by the *one* favored mode of presentation, because there is not such a unique favored perspective. The following example supplies evidence for this point.

the soccer game Suppose you are attending a soccer game. All of the 22 players are in your perceptual field. You know their names, say a, b, c, ..., but you don't recognize any of them. Consider the following sentence:[20]

(142) a. Anyone might be anyone.

b. $\forall x\forall y\Diamond(x = y)$

It seems to me that (142) can be uttered in this situation. However, if we assume (moderate) slicing, (142) is inconsistent. The source of the difficulty is the uniqueness presupposition behind the total interpretation of variables. Intensional properties such as 'possibly being anyone' are not traits of individuals *simpliciter*, but depend on the perspective under which these individuals are looked at. Examples like (142) show that there is not one direct way of looking at the universe of discourse that characterizes the domain of quantification once and for all, instead different perspectives seem to supply different sets of ultimately partial objects over which we can quantify.

[19] See the classical articles Quine (1956), Kaplan (1969) and Kripke (1979) and more recently Gerbrandy (2000). See also the previous two chapters of this thesis.

[20] This is a modification of an example of Paul Dekker.

Synopsis The diagram below summarizes the contents of this section:

	RA	**SL**	**MS**
quant. var.	partial $\rightsquigarrow$ undersp 1	total $\rightsquigarrow$ oversp 2	total $\rightsquigarrow$ oversp 2
free var.	partial $\rightsquigarrow$ undersp 2	total $\rightsquigarrow$ oversp 1	partial $\rightsquigarrow$ undersp 2

3.4 Dynamic Quantification under Cover

In order to overcome the problems of over- and underspecification, I propose a new style of dynamic quantification that lies between random assignment and slicing, and which treats free and quantified variables in a uniform way. On the one hand, as in slicing, free and quantified variables range over alternative definite elements of some domain, and with respect to the different choices of these elements the interpretation proceeds in a parallel fashion. In this way, variables vary enough to avoid the underspecification problems. On the other hand, the overspecification problems are solved by allowing not one but many ways of conceiving the individuals over which we quantify. Different sets of possibly non-rigid concepts that cover the whole universe and that do not consider any individual more than once constitute suitable candidates for the domain of quantification.

Definite Subjects

In dynamic semantics, two levels of objects are assumed: the individual elements of the universe of discourse, and the partial entities that constitute the interpretations of variables in information states. The latter are introduced as items in conversation and can change, for instance by growing less partial, as the conversation proceeds. As we saw,[21] in Dekker (1993), these entities are called partial objects and are defined as functions that assign to each possibility in a state the value of the corresponding variable in that possibility. I extend Dekker's definition of partial objects and call a *subject* in an information state any mapping from the possibilities (world-assignment pairs) in the state to the individuals in the universe of discourse. Notice that in addition to explicitly introduced discourse items, potential items also count as subjects in a state.

Among the subjects, we can distinguish *rigid subjects* and *(in)definite subjects.* Rigid subjects are the constant functions among the subjects. Definite subjects are those that assign the same value to all possibilities that have the same world

[21] See footnote 2.

parameter. Definite subjects are contextually restricted (individual) concepts. They are *definite* in that they have a single value relative to a single world, but *partial* in that they may have different values relative to different worlds, and hence they do not always determine one individual. Indefinite subjects are subjects that are not definite, i.e., those assigning different values to possibilities with the same factual content.

In RA and MS, the presence of indefinite subjects as interpretation of some discourse items in a state reflects the indeterminacy of the addressee's perspective. Note that from the speaker's point of view indefinite items are senseless. Consider the following dialogue discussed in Dekker (1997):[22]

K: Yesterday a man came into my office who inquired after the secretary's office.
J: Was he wearing a purple jogging suit?
K: If it was Arnold, he was, and if it was somebody else, he was not.

Consider the following context: (α) K knows that Arnold *and* somebody else went to his office inquiring after the secretary's office. Dekker observes that in such a context, K's reply is odd, because K should have made up his mind about whom he wanted to talk before starting to tell the story. But now imagine another scenario: (β) K, who is blind but knows that Arnold always wears eccentric jogging suits, was wondering from the beginning whether it was Arnold who went to his office *or* somebody else. In β, the dialogue becomes quite natural.

Speakers do not introduce indefinite subjects (in scenario α the dialogue is odd), but may introduce non-rigid subjects (in scenario β the dialogue is natural). It seems fair to conclude that speakers introduce definite subjects. Now, dynamic semantics models the addressee's updating procedure and addressees often lack information about which definite subjects speakers intend to refer to. Questions

[22] See Dekker (1997) and van Rooy (1997). In (1997), Dekker uses such dialogues as a motivation for his interesting notion of dynamic support defined in term of *links* (see also Dekker (2000b)). As I said in a previous footnote, the adoption of Dekker's compositional support also allows a solution to the underspecification 2 problems. In van Rooy (1997), chapter 2, such dialogues are accounted for by assuming that specific indefinite NPs introduce speaker's referents, which are definite objects, rather than underspecified discourse items. Note that on van Rooy's account, overspecification 1 is also avoided, by positing two kinds of pronouns, referential and descriptive, the former referring back to total speaker's referents, the latter denoting possibly partial, but definite objects. Van Rooy's interesting distinction between referential and descriptive pronouns allows him to account for Barbara Partee's famous bathroom examples in an enlightening way. However, as I argued above, overspecification is not restricted to free variables, but arises for syntactically bound variables as well, and van Rooy's analysis does not have an explanation of the latter cases. The analysis I propose in the following sections avoids overspecification in general, but does not apply to the bathroom cases since the pronouns there are not dynamically bound. A combined approach might be the correct one, which treats syntactically and dynamically bound pronouns as ranging over elements of conceptual covers and dynamically unbound pronouns as van Rooy's descriptive pronouns.

like 'Who do you mean?' or 'Who are you talking about?' represent states of ignorance of this kind. In RA and MS, this type of ignorance is modeled in the same way as ordinary ignorance about what is the case, that is, by the presence in the information state of a number of world-assignment pairs in which the different individuals that the speaker might have in mind are represented by the different values that the relevant variable can take. A consequence of this strategy is the presence in states of indefinite subjects. In SL, instead, (lack of) information about the speaker's intentions is modeled on a higher level, namely by the presence of different alternative updates that run in a parallel fashion. In SL, the possible speaker's referents are modeled by the rigid subjects that constitute the interpretation of the relevant variable in the alternative parallel states. Going back to the dialogue above, K's reply is intuitively acceptable just in case K is assumed to be in doubt about the identity of the visiting person(s), so in a context like β. In cases in which K is assumed to know who the relevant persons were, the sentence is clearly unacceptable, because speakers cannot have doubts about their own intentions (as shown by the markedness of the sentence in α). Now, it is not obvious how RA, MS or SI can account for these intuitions. If the item introduced by K is modeled by an indefinite subject (RA and MS), the sentence would be judged acceptable even in case K is assumed to be omniscient with respect to information about the world. On the other hand, if we avoid indefinite subjects, but admit only rigid subjects (SL), then K's reply is never judged acceptable, which is also incorrect. The solution which suggests itself is to rule out indefinite subjects, while allowing possibly non-rigid ones and this is also my proposal.

	x
w_1	a
w_1	b
w_2	a
w_2	b

1. partial

	x
w_1	a
w_2	a

	x
w_1	b
w_2	b

2. total

	x
w_1	a
w_2	a

	x
w_1	a
w_2	b

	x
w_1	b
w_2	a

	x
w_1	b
w_2	b

3. definite

I propose to let variables range over definite subjects.[23] The interpretation

[23]Semantically my approach thus comes close in spirit to alternative approaches to anaphora

of the existential quantifier will involve splitting up a state as in the slicing procedure. The definite subjects (possibly non-rigid ones) which the speaker might have in mind are considered one by one as disjoint alternatives. It seems that in this way we can avoid underspecification without falling into overspecification. Since variables are taken to range over alternative elements of some domain, we avoid Dekker's or Heim's problem. In addition, since they can vary over non-rigid subjects, we have a good hope of solving the overspecification problems as well. Definite subjects seem to be the 'something in between' that we were looking for. Quantification over concepts is, however, quite an intricate affair. Difficulties arise almost immediately from the fact, evident from the picture above, that there are strictly more concepts in a state than individuals in the universe of discourse. Consider the following two examples.

the winner If we let quantifiers range over the set of all definite subjects, the following is a valid scheme:

(143) $\forall x \Diamond \phi \rightarrow \Diamond \forall x \phi$

This is clearly undesirable.[24] Suppose a game has been played; (143) says that if it is known that there are some losers ($\neg \Diamond \forall x W x$), but we have no clue about who won, then we have no way of expressing this ignorance in a single quantified statement (since $\neg \forall x \Diamond W x$ must also be true). (See also the shortest spy problems emerging for the system CIA, discussed in chapter 2, section 2.3.1.) Another example showing the same point is the smallest flea case.

the smallest flea Consider the following two sentences:

(144) Any flea might be the smallest flea.

(145) The biggest flea might be the smallest flea.

If we quantify over all concepts, a generalized version of universal instantiation holds and we can derive (145) from (144). This means that in ordinary situations in which fleas differ in size, (144) is never accepted. There will always be an element in the quantificational domain that falsifies it, for instance *the biggest flea*. Thus ignorance about the smallest flea's identity is inexpressible in such situations.

The examples above seem to show that quantifiers in natural language do not range over representations of individuals without further restrictions. If sentences

like e.g. E-type theory approaches as those of Evans (1977), Neale (1993), and in particular Slater (1986).

[24] Quine, though discussing a different point, shows the implausibility of the equivalent scheme $\Box \exists x \phi \rightarrow \exists x \Box \phi$: '...in a game of a type admitting of no tie it is necessary that some one of the players will win, but there is no one player of whom it may be said to be necessary that he win.' Quine (1953), p. 148.

like (144) were to quantify over representations, then we would have to accept the derivation of (145) from (144) as a trivial one. The fact that instead this conclusion strikes us as counter-intuitive means that natural language quantifiers do not work in this way. When we talk, we talk about individuals, not about representations of individuals, even in situations in which we lack information about their identity or are misinformed about that. To capture this feature of natural language quantifiers, we need a notion of *aboutness* which can work in situations of *partial information*. The traditional characterization of aboutness in terms of rigidity, implicit in (moderate) slicing, is inadequate in these cases. As we saw, in situations of partial information, we do not quantify over total objects (because we cannot). However, to deny the claim that quantifiers range over individuals in a direct way, we need not assume that we quantify *over* representations - it is enough to say that we quantify over individuals, but *under* a representation. Natural language quantifiers range over individuals under a *perspective*. To give some content to this abstract claim, let's consider the following example[25] in which we see such perspectives at work.

the butler Suppose a butler and a gardener are sitting in a room. One is called Alfred and the other Bill. We don't know who is who. In addition, assume that the butler committed a terrible crime. Consider now the following two discourses:

(146) The gardener didn't do it. So it is not true that anybody (in the room) might be the culprit.

(147) Alfred might be the culprit. Bill might be the culprit. So anybody (in the room) might be the culprit.

It seems to me that both (146) and (147) can be uttered in such a situation given the right circumstances. We can intuitively explain what is going on as follows: intensional properties, such as perhaps being the culprit, do not properly apply to individuals *simpliciter*, but depend on the perspective under which these individuals are conceived. Although the universal quantifier ranges over the same set in the two discourses, namely the set containing the two people in the room, in the two cases, the two individuals are identified from two different angles. For this reason no contradiction arises. In (146), individuals are looked at under the perspective of their profession; in (147) they are identified as referents of some proper name. Under the latter identification method, *the butler* may be Alfred or may be Bill. Yet, if we assume the other perspective, we can think of *the butler* as standing for a single object contrasted with *the gardener*. Perspectives are determined by contextual factors. In these two specific cases, the relevant contextual information is supplied by the preceding sentences, which, by mentioning one concept or the other, suggest one or the other way of classifying the domain.

[25] For more about examples of this kind see Gerbrandy (2000).

A natural way of representing a perspective over the universe of discourse is by means of a set of concepts. However, not all collections of concepts will do. The set of *all* concepts, for instance, is not a good candidate, as is evident from the winner and the smallest flea examples above. But there are many more inadequate conceptualizations.

Take a situation similar to the one above. Again we have Alfred and Bill sitting in some room, we know that one of the two is the butler, the other is the gardener, but we don't know who is who. Suppose you are interested in determining whether for anyone in the room, it is consistent with your information that an arbitrary property, say being bald, holds.

(148) Anyone might be bald.

Which of the following sentences constitutes a sufficient ground for a correct assertion of (148)?

(149) Alfred might be bald and Bill might be bald.

(150) The gardener might be bald and the butler might be bald.

(151) Alfred might be bald and the butler might be bald.

(152) Bill might be bald and the gardener might be bald.

In this particular situation, only the first two can ground (148). A derivation of (148) from either (151) or (152), would not be accepted as an example of correct reasoning. Even if the context may raise them as sets of salient concepts, the set consisting of *Alfred* and *the butler* as well as the set consisting of *Bill* and *the gardener* are not good conceptualizations in this specific case.[26] The reason for this is that, intuitively, they do not provide a uniform perspective over the universe of discourse: they mix up different perspectives and they do not cover the domain of individuals in an exhaustive way. In the following section I propose a way to formalize these intuitions.

Conceptual Covers

In this section, I restate the definition of the notion of a conceptual cover, introduced in the previous chapters, which is shown to be needed to account for the issues discussed in the previous sections.

Given a set of worlds W and a set of individuals D, an individual concept is any total function from W to D. A conceptual cover is a set of individual concepts

[26] Their inadequacy doesn't follow from the fact that they use definite descriptions *and* proper names, but depends on the specific information supported in this case. In other situations, such sets can provide good conceptualizations.

that satisfies the following condition: in a conceptual cover, in each world, each individual constitutes the value of one and only one concept.

Given a set of possible worlds W and a universe of individuals D, a *conceptual cover* CC based on (W, D) is a set of functions $W \to D$ such that:

$$\forall w \in W: \ \forall d \in D: \ \exists! c \in CC : c(w) = d$$

Conceptual covers are sets of concepts which exhaustively and exclusively cover the domain of individuals. In a conceptual cover, each individual d is 'seen' by at least one concept in each world (*existence*), but in no world is an individual counted more than once (*uniqueness*).

Two typical examples of conceptual covers are the following (let $\mathcal{C}$ be the set of individual constants in $\mathcal{L}$):

1. $RC = \{\lambda w \ d \mid d \in D\}$ (rigid cover)
2. $NC = \{\lambda w \ w(a) \mid a \in \mathcal{C}\}$ (naming)

RC is the set of constant concepts and NC is the set of concepts that assign to every world the denotation of a certain individual constant in that world (assuming exhaustive and exclusive naming practices). These two covers can be taken to model the two identification methods that played a role in most of the examples above (cf. for instance (142)), namely identification by ostension and identification by naming. However, these are just two among the many methods of identification that we normally assume when we think or talk about objects in our everyday practices. Other families of identification methods are, for instance, identification by description as in example (146), or by recognition like in the cases in which we identify strangers by bringing to mind the visual image of their faces that we perceived at one time. Our theory has no problem providing enough conceptual covers to model this multitude of identification methods.[27]

I propose to let variables range over the elements of a contextually supplied conceptual cover. The existential and the universal quantifiers will behave as ordinary quantifiers, that is, even if, technically, they range over concepts, the effect obtained is that of quantification over genuine individuals. It is precisely this ordinary type of quantification that motivates the two constraints on conceptual covers specified above, in particular the uniqueness condition which serves to guarantee that the objects over which we quantify eventually correspond to determinate individuals that can be said to be identical with themselves and distinct from one another. Sets of overlapping concepts do not characterize sets of genuine individuals in this sense. Consider again the situation described in the butler example above. Alfred and Bill are sitting in some room, we know that one of the two is the butler and the other is the gardener, but we don't know who

[27] See chapter 4, proposition 4.1.3.

is who. Take the set A consisting of the concepts *Alfred* and *the butler*. First of all, observe that A is not a conceptual cover. Given our assumptions, there will be some world w, in which someone is counted twice (namely the individual which is Alfred and the butler in w), and someone else is not 'seen' at all (namely the individual which is Bill and the gardener in w). So A does not satisfy either the uniqueness or the existence condition. Now, given the situation, the two elements of A cannot be regarded as standing for two determined individuals. Since we wonder whether Alfred is the butler, *Alfred* and *the butler* might be one individual or two. Crucially, inside exhaustive and exclusive sets of concepts, this kind of indeterminacy does not arise. Consider the set B consisting of *the butler* and *the gardener*, which, given our assumptions, is a conceptual cover. When taken in combination with *Alfred*, the concept *the butler* gives rise to individuation problems, but here, contrasted with *the gardener*, it comprises a completely determinate individual. Thus, only as an element of B and not as an element of A, *the butler* is capable of serving as a value of some bound variable.

To conclude, the elements of a conceptual cover represent the entities we quantify over and that we experience only via one or the other mode of presentation; yet it would be misleading to identify them with these modes of presentation.[28] The elements of a conceptualization are the individuals themselves, just thought, conceived or identified in a particular way.

Quantification under Conceptual Covers

To define quantification under conceptual covers, I need an operation that extends information states in the appropriate way.

3.4.1. DEFINITION. [c-extensions] Let $\sigma \subseteq I^X, x \notin X$ & $c \in D^W$. Then

$$\sigma[x/c] = \{i[x/c(w_i)] \mid i \in \sigma\}$$

C-extensions lie between global and individual extensions. They introduce fresh variables and interpret them as certain definite subjects. Dynamic quantifiers are defined in terms of c-extensions; they range over elements of a contextually-given conceptual cover and, only indirectly, over the individuals in the universe. In this way, quantification is relativized to a particular way of conceptualizing the domain.

I add a special index $n \in N$ to the variables in $\mathcal{L}$. These indices range over conceptual covers and their value is assumed to be pragmatically supplied. As in chapter 2, I will write $\mathcal{V}_n$ to denote the set of variables indexed with n and $\mathcal{V}_N$ to denote the set $\bigcup_{n \in N}(V_n)$.

A model for this richer language $\mathcal{L}_{CC}$ is a triple $\langle D, W, C\rangle$ where D and W are as above and C is a set of conceptual covers based on (W, D). The interpretation

[28] Or with Fregean senses, characterized as ways of thinking of the referent of some singular term.

function $[\bullet]$ is relativized to *conceptual perspectives* $\wp$ which are functions from N to C.[29] Only the interpretation of dynamic quantifiers is directly affected by this relativization.

3.4.2. Definition. [CC-Quantification]

$$\sigma[\exists x_n \phi]^{\wp}_{CC}\sigma' \quad \text{iff} \quad \sigma[x_n/c][\phi]^{\wp}_{CC}\,\sigma' \text{ for some } c \in \wp(n)$$

A quantifier Qx_n, evaluated under a conceptual perspective $\wp$, is taken to range over the conceptual cover assigned by $\wp$ to n. The fact that each variable occurs with its own index allows different occurrences of quantifiers to range over different sets of concepts. Although different quantifiers range over the same sort of individuals, these may be identified in different ways.

If we assume quantification under cover, slicing and the classical theory of quantification arise as a special case, namely when all indices are assigned the rigid cover. Let $\wp^*(n) = RC$, for all n. We then obtain:

$$\sigma[\exists x_n \phi]_{SL}\sigma' \quad \text{iff} \quad \sigma[\exists x_n \phi]^{\wp^*}_{CC}\,\sigma'$$

RA and MS can be defined as derived notions in terms of c-extensions.

$$\begin{array}{lll} \sigma[\exists x_n \phi]_{RA}\sigma' & \text{iff} & (\cup_{c\in RC}\{\sigma[x_n/c]\})[\phi]_{RA}\sigma'; \\ \sigma[\exists x_n \phi]_{MS}\sigma' & \text{iff} & \sigma' = \cup_{c\in RC}\{\sigma'' \mid \sigma[x_n/c][\phi]_{MS}\sigma''\}. \end{array}$$

We can now relativize the notion of support to pragmatic contexts in an obvious way.

3.4.3. Definition. [CC-Support] Let $\wp$ be a conceptual perspective, σ be in Σ_M, and ϕ in $\mathcal{L}_{CC}$.

$$\sigma \approx\!\!|_{\wp}\, \phi \quad \text{iff} \quad \exists\sigma' : \sigma[\phi]^{\wp}\sigma' \;\&\; \sigma \prec \sigma'$$

A state σ supports a sentence ϕ under a perspective $\wp$ iff all possibilities in σ survive simultaneously in at least one of the states resulting from updating σ with ϕ under $\wp$.

3.5 Applications

In this section, I show how the use of conceptual covers solves the problems discussed earlier in this chapter.

[29] See chapter 1.

Underspecification Since variables in quantified contexts are taken to range over alternative definite objects, underspecification **1** is avoided. I will first consider Dekker's problem, which concerns the following two sentences:

(153) a. Someone might be the culprit.

b. $\exists x_n \Diamond P x_n$

(154) a. Someone is certainly not the culprit.

b. $\exists x_n \neg \Diamond P x_n$

Sentences (153) and (154) do not contradict each other, because different definite subjects are considered in isolation and they are not absorbed into a single indefinite one (as in *RA*). For instance, let σ be a state and c_1, c_2 be two concepts in some conceptual cover CC such that only c_1 takes as values individuals that have the property P in some possibilities of σ. Such a state σ will support both (153) and (154) under a perspective $\wp$ which assigns CC to n, since $\sigma[x_n/c_1] \mathrel{|}\!\!\approx_{\wp} \Diamond P x_n$ and $\sigma[x_n/c_2] \mathrel{|}\!\!\approx_{\wp} \neg \Diamond P x_n$. Heim's problem can be handled analogously:

(155) a. A fat man was pushing his bicycle.

b. $\exists x_n[\text{fat-man}(x_n) \wedge \partial(\exists y_m \text{ bike-of}(x_n, y_m)) \wedge \text{ pushing}(x_n, y_m)]$

If we assume quantification under cover, examples like (155) can be taken to project an existential presupposition, rather than a universal one, as we intuitively expect. As an illustration, consider the following variation of the butler situation. We have two fat men, Alfred and Bill. One is a gardener, the other is a butler. We don't know who is who, but it is established that while Alfred has a bike, Bill has none. As we saw, in *RA*, (155) is undefined in such a situation, because not all fat men have a bike:

$$\textbf{RA :} \quad \begin{array}{|c|}\hline w_1 \\ \hline w_2 \\ \hline \end{array} \quad [\exists x \text{ fat-man}(x)]_{RA} \quad \begin{array}{|c|c|}\hline & \mathbf{x} \\ \hline w_1 & a \\ \hline w_1 & b \\ \hline w_2 & a \\ \hline w_2 & b \\ \hline \end{array} \quad [\partial(\exists y \text{ bike-of}(x, y))]_{RA} \qquad (\ddot{\frown})$$

In *CC*, instead, if the index n is assigned a cover containing the concepts *Alfred* and *Bill*, (155) is defined, since the two possibilities of x_n being Alfred or Bill are considered in isolation (in the picture such a cover is assumed to be the rigid cover):

$$\mathbf{CC:}\ \begin{array}{|c|}\hline w_1\\ \hline w_2\\ \hline\end{array}\ [\exists x_n \text{ fat-man}(x_n)]^{\wp}_{CC} \left\langle \begin{array}{l} \begin{array}{|c|c|}\hline & \mathbf{x_n}\\ \hline w_1 & a\\ \hline w_2 & a\\ \hline\end{array}\ [\partial(\exists y_m \text{ bike-of}(x_n, y_m))]^{\wp}_{CC}\quad (\ddot\smile) \\[2em] \begin{array}{|c|c|}\hline & \mathbf{x_n}\\ \hline w_1 & b\\ \hline w_2 & b\\ \hline\end{array}\ [\partial(\exists y_m \text{ bike-of}(x_n, y_m))]^{\wp}_{CC}\quad (\ddot\frown) \end{array}\right.$$

Note that, the sentence is still undefined, if n is assigned a cover containing the concepts *the butler* and *the gardener*, but not because a universal presupposition is projected in this case, but because, under such a cover, there is no possible intended fat man about whom it is established that he has a bike.

Underspecification **2** is also avoided, since only concepts may be introduced as new items. I just illustrate how the wrong suspect case is handled in the new system. Intuitively, example (156) comes out incoherent because there are no possible concepts under any conceptualization that can satisfy the two conjuncts at the same time.

(156) a. Someone might be the culprit. She is not the culprit.

b. $\exists x_n \Diamond P x_n \wedge \neg P x_n$

Formally, the incoherence of (156) follows from the fact that it is impossible for a state resulting from a successful update with $\exists x_n \Diamond \phi$ to support $\neg\phi$, in combination with the 'compositionality' of support. In the present semantics, if a conjunction $\phi \wedge \psi$ is supported in a state σ, then ψ must be supported in some σ' resulting from a successful update of σ with ϕ.

3.5.1. Fact. [Support of Conjuncts] Let $\sigma \in \Sigma_M$ and ϕ, ψ in $\mathcal{L}$.

$$\sigma \mathrel{|\!\approx_{\wp}} \phi \wedge \psi \;\;\Rightarrow\;\; \sigma \mathrel{|\!\approx_{\wp}} \phi \;\;\&\;\; \exists \sigma' : \sigma[\phi]^{\wp}\sigma' \;\;\&\;\; \sigma' \mathrel{|\!\approx_{\wp}} \psi$$

Crucial to the proof of this fact[30] is the following property of the new system:

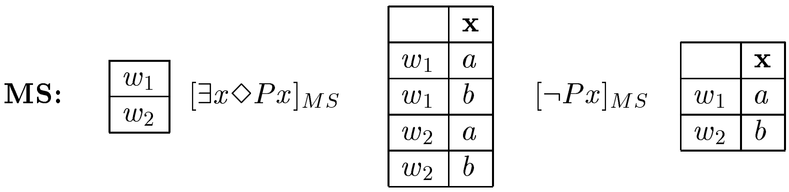
3.5.2. Fact. [Unique Extension] Let $\sigma[\phi]^{\wp}\tau$ for some perspective $\wp$, some states $\sigma, \tau \in \Sigma_M$ and some $\phi \in \mathcal{L}$, then the following holds:

$$\forall i \in \sigma : \forall j_1, j_2 \in \tau : i \prec j_1 \;\&\; i \prec j_2 \rightarrow j_1 = j_2$$

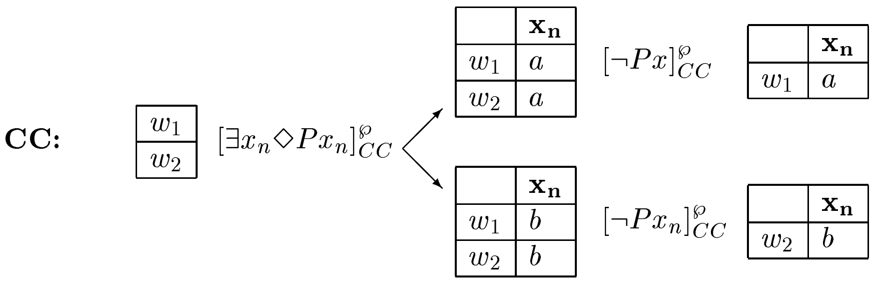
For any update in a system satisfying such a property, no two possibilities in the output state can extend one and the same possibility of the input state. Typical examples of systems in which the unique extension property does not hold are RA and MS, namely systems that allow a partial interpretation for free variables.

[30] We will come back to this later.

As an intuitive illustration of how the wrong suspect problem is solved, I will compare the interpretation procedures for (156) in MS, which allows branching of possibilities, and in the new system, in which the unique extension property holds. Let the input state be the state σ^* as above, consisting of two possibilities that supports the information that either individual a (in w_2) or individual b (in w_1) is P. As we saw, in MS, that allows a partial interpretation for free variables, σ^* supports the sequence and hence the latter is predicted to be coherent.



If, on the other hand, we adopt the style of quantification that I am proposing, we avoid this problem: the two initial possibilities do not survive both in one of the output states under any conceptual cover. In the picture below, we consider as an illustration the case in which $\wp(n) = RC$.



As a matter of fact, no state can be found that supports (156) under any conceptualization. The sentence is incoherent.

Overspecification Since variables are not taken to range over individuals *simpliciter*, but over individual under a conceptualization, the overspecification problems are solved as well. The inequalitarian attitude towards ways of identifying objects implicit in (moderate) slicing is overcome and different identification methods are given equal status. We can look at the individuals in the universe under different perspectives and, if the context justifies it, we can change perspective within the same discourse. Problems of identification can be represented as problems of mapping elements from different conceptualizations onto each other. As a result, the overspecification **1** cases are solved.

(157) a. Someone did it. It might be anyone.

b. $\exists x_n P(x_n) \wedge \forall y_m \Diamond (x_n = y_m)$

Examples like (157) come out coherent, if interpreted in a perspective $\wp$ that assigns different conceptual covers to n and m. For example, if the existential quantifier introduces non-rigid subjects and the universal ranges over the rigid conceptualization ($\wp(m) = RC$), a state like σ^* above supports (157).

Overspecification **2** is avoided in a similar way. Example (158) can be accepted, if x_n is not taken to range over the cover representing demonstrative identification, RC.

(158) $\neg \exists x_n \Box (x_n = \text{this})$

A good example of this obtains when n is assigned naming. Thus we can express ignorance about the identity of some object of perception, and in addition, by shifting conceptualization we can account for any situation of partial identification in an enlightening way. Examples like 'I wonder who Alfred is', or 'I wonder who the culprit is' are not problematic for this approach.[31] The soccer game case is explained as well.

(159) a. Anyone might be anyone.

b. $\forall x_n \forall y_m \Diamond (x_n = y_m)$

Example (159) is acceptable, if x_n and y_m are taken to range over different conceptualizations. In this specific situation, n and m are assigned as value the cover by perception and the cover by names respectively. Notice, however, that if no shift of cover is assumed the sentence is unacceptable: $\forall x_n \forall y_n \Diamond (x_n = y_n)$ is inconsistent in the present approach (unless the domain contains a single individual).

Cardinality Since the set of all concepts is not a conceptual cover, the winner problem and the smallest flea problem do not occur. The problematic scheme:

(160) $\forall x_n \Diamond \phi \rightarrow \Diamond \forall x_n \phi$

is *not* valid and hence sentences like 'Anyone might be the winner' can be accepted in situations in which it is known that there are some losers. Furthermore, only a restricted version of universal instantiation holds:

(161) $(\forall x_n \phi \wedge \exists y_n \Box (t = y_n)) \rightarrow \phi[x_n/t]$

So, since the universal sentence 'Any flea might be the smallest flea' can be accepted only under a conceptualization that does not contain *the biggest flea*,[32] the problematic implication to 'The biggest flea might be the smallest flea' is blocked.

[31] See chapter 1.

[32] Unless we have a domain with a single flea.

To summarize, if we assume quantification under cover, underspecification does not occur since only definite subjects may constitute interpretations of variables. At the same time, overspecification is also avoided since different occurrences of quantifiers may range over different sets of (possibly partial) concepts. Finally, by taking as domain of quantification only sets of concepts which exhaustively and exclusively cover the universe of individuals, we avoid the cardinality problems that normally arise when we quantify over concepts rather than objects.

3.6 Conceptual Covers in Dynamic Semantics

In the previous sections, I have compared three different intensional dynamic systems using three different styles of quantification: RA, MS and SL, and I have discussed specific problems arising for each of them. I have then introduced a new style of dynamic quantification, quantification under cover, and I have shown how, by adopting such a form of quantification, the previously discussed problems are avoided. In this section, I show how dynamic semantics under conceptual cover can be formulated in a much more elegant fashion.

In what follows, I present Dynamic Modal Predicate Logic under Conceptual Covers (CC). I abandon the relational version of MDPL, which I used for the comparison of the four different styles of dynamic quantification, and I formulate the CC system in a style which is closer to the logic presented in chapter 2. The proposed semantics will be an 'under cover' version of the Dynamic Modal Predicate Logic (DMPL) introduced in van Eijck and Cepparello (1994). The choice of formulating the semantics in the style of DMPL is not unmotivated. Indeed, this formulation clearly shows important properties of CC-quantification, because it crucially exploits them. Furthermore, DMPL has a distinct advantage over MDPL, namely that the reuse of variables does not cause the 'downdate' problem (see footnote 11). The fact that quantification under cover allows this kind of re-formulation constitutes a further motivation for its assumption – for RA and MS cannot be reformulated in this way. I start by discussing the issue of non-accessibility in an MDPL style formulation of CC-quantification.

Non-accessible states

We can think of meanings in a system S as describing transitions between different information states. I will call a state *S-accessible*, if it is reachable from the state of minimal information $1 = \{\langle w, \emptyset\rangle \mid w \in W\}$ by a number of S-transitions. I will call a state *CC-accessible under* $\wp$, if it is reachable from the state of minimal information by a number of CC-transitions under $\wp$.

CC-accessible states under some perspective $\wp$ satisfy the following condition: in such a state, variables which are indexed in a uniform way are interpreted as elements of one and the same conceptual cover. I will call this condition $\wp$-

uniformity. Let $[x_n]_\sigma$ denote the partial object which constitutes the interpretation of the variable x_n in σ, i.e., $[x_n]_\sigma$ is the function $f : \sigma \to D$ such that $\forall i \in \sigma : f(i) = i(x_n)$; let c_σ denote the restriction of the concept c to σ, i.e., c_σ is the function $f : \sigma \to D$ such that $\forall i \in \sigma : f(i) = c(w)$. The following proposition trivially holds:

3.6.1. Proposition. [$\wp$-uniformity] Let $\sigma \subseteq I^X$ be a CC-accessible state under some perspective $\wp$. Then the following holds:

$$\forall n \in N : \forall x_n \in X : \exists c \in \wp(n) : [x_n]_\sigma = c_\sigma$$

As a corollary of this proposition, CC-accessible states have the important property that they do not contain indefinite subjects.

3.6.2. Corollary. [Definiteness] Let σ be a CC-connected state. Then the following holds:

$$\forall i, j \in \sigma : w_i = w_j \;\Rightarrow\; i = j$$

Since not all MDPL information states are $\wp$-uniform or definite, it follows that, in CC, not all states are accessible.[33] In what follows I want to argue that, firstly, non-accessible states are not useful, and, secondly, they can be harmful and hence we have good reasons to get rid of them.

There are many ways in which we can acquire new information, and linguistic communication is just one of these ways. We may wonder whether non-accessible states might be useful to encode information, obtainable by some non-linguistic means, which cannot be encoded by some accessible state. I do not believe this is the case. The only kind of information which can be encoded in a non-accessible state and cannot be encoded in an accessible state concerns indefinite subjects. In ordinary dynamic semantics indefinite subjects originate from the interpretation of indefinite NPs, but I have already argued in favor of a definite interpretation of such expressions. Indefinite subjects are used to express lack of information about the actual intended denotation of a discourse item,[34] but, as we have seen, this kind of ignorance can also be expressed by means of (a set of) definite state(s).[35] Thus, non-accessible states do not add any expressive power to our system, and, therefore, we do not have any reason to maintain them. Instead, we have reasons to eliminate them. Indeed, non-accessible states affect the logical notions

[33]Something similar holds for SL as well. Note that in RA and MS all states are accessible.

[34]For discussion about the role of discourse referents see Dekker (2000a) and Zimmermann (1999).

[35]This point can be compared with that of Stalnaker in the debate with Lewis about indexical belief. See Lewis (1979) and Stalnaker (1981).

of our system in an undesirable way. In order to see this, consider the notion of coherence, which we have focused upon in the previous sections. Compare the following two notions of coherence where the first involves existential quantification over all states and the second restricts quantification to accessible states (I use $\mathrm{Acc}_{\wp}(\sigma)$ to denote that σ is CC-accessible under $\wp$):

3.6.3. Definition. [CC-Coherence 1] Let ϕ be in $\mathcal{L}_{CC}$.

$$\approx\!\!\!\!|_{CC1}\, \phi \text{ iff } \exists M, \exists \wp, \exists \sigma \in \Sigma_M : \sigma \neq \emptyset \;\&\; \sigma \models\!\!\!\!\approx_{\wp} \phi$$

3.6.4. Definition. [CC-Coherence 2] Let ϕ be in $\mathcal{L}_{CC}$.

$$\approx\!\!\!\!|_{CC2}\, \phi \text{ iff } \exists M, \exists \wp, \exists \sigma \in \Sigma_M : \mathrm{Acc}_{\wp}(\sigma) \;\&\; \sigma \neq \emptyset \;\&\; \sigma \models\!\!\!\!\approx_{\wp} \phi$$

Coherence 1 and 2 are not the same notion. Consider the following sentence:

(162) $\Diamond\, x_n = y_n \wedge \Diamond\, x_n \neq y_n$

If we assume the second notion of coherence, (162) is incoherent. If we assume the first, it is not. Indeed, a non-accessible state which does not satisfy $\wp$-uniformity can support the sentence:

(163) $\approx\!\!\!\!|_{CC1} \Diamond\, x_n = y_n \wedge \Diamond\, x_n \neq y_n$

(164) $\not\approx\!\!\!\!|_{CC2} \Diamond\, x_n = y_n \wedge \Diamond\, x_n \neq y_n$

Example (162) expresses the negation of the uniqueness condition on conceptual covers. Since x_n and y_n are equally indexed, we expect them to range over elements of the same cover, but their denotations are neither separated (they coincide in some possibility) nor equal (do not coincide everywhere). Above I have argued that it is because of the uniqueness condition that covers constitute suitable domains of quantification. Since the elements of a cover do not merge and split once we move from one possibility to the other, they behave like genuine individuals and, therefore, are capable of serving as the value of some bound variable. For this reason, we would like a sentence like (162), which negates the very nature of conceptual covers, to be incoherent in CC. The most natural way to obtain this consists in eliminating non-accessible states. We can model CC-quantification in an MDPL style dynamic semantics, in such a way that non-accessible states are neutralized. We can restrict all logical notions to accessible states or redefine the notion of a state in such a way that only $\wp$-uniform states can count as states. None of these moves is particularly elegant. I take this as an indication that an MDPL style dynamic semantics, which was the right framework for the comparison among the four styles of dynamic quantification, is not a natural choice for CC (and SL). In what follows, building on van Eijck and Cepparello (1994), I reformulate quantification under cover in a way which does better justice to its arguably desirable features. In the new formulation, the problematic non-accessible states are not even definable. Furthermore, in the new semantics, as in the original Groenendijk and Stokhof formulation of Dynamic Predicate Logic, 'downdates' are not problematic. Thus, all the complications connected to the issue of the reuse of variables which arise for an MDPL style formulation are avoided here.

Dynamic Modal Predicate Logic under Cover

A model for a language $\mathcal{L}_{CC}$ is, as above, a triple $\langle W, D, C\rangle$ where C is a set of conceptual covers on D and W. A *state* $s \in S_M$ in the new formulation is a subset of W, as in the original formulation of update semantics in Veltman (1997). A CC-*assignment* g is a function mapping conceptual cover indices $n \in N$ to conceptual cover elements of C and n-indexed individual variables $x_n \in \mathcal{V}_n$ to concepts which are elements of $g(n)$ (see definition 2.4.1, in chapter 2).[36] Thus, the information that, in the previous formulation was encoded separately in the old assignments and in the conceptual perspectives, is encoded in an integrated fashion here. Terms are interpreted as follows:

3.6.5. Definition. [CC-Interpretation of terms]

(i) $[t]_{w,g} = g(t)(w)$, if t is a variable;

(ii) $[t]_{w,g} = w(t)$, if t is a constant.

Sentences ϕ are interpreted with respect to pairs of input-output assignment functions $[\phi]^g_h$. The denotation $[\phi]^g_h$ of ϕ with respect to g and h is a function from states to states.

3.6.6. Definition. [CC-Semantics]

$$
\begin{array}{rll}
s[Rt_1, ..., t_n]^g_h = t & \text{iff} & g = h \ \& \ t = \{w \in s \mid \langle [t_1]_{w,g}, ..., [t_n]_{w,g}\rangle \in w(R)\}; \\
s[\neg\phi]^g_h = t & \text{iff} & g = h \ \& \ t = \{w \in s \mid \neg\exists k : w \in s[\phi]^g_k\}; \\
s[\exists x_n]^g_h = t & \text{iff} & g[x_n]h \ \& \ t = s; \\
s[\phi \wedge \psi]^g_h = t & \text{iff} & \exists k : (s[\phi]^g_k)[\psi]^k_h = t; \\
s[\Diamond\phi]^g_h = t & \text{iff} & g = h \ \& \ t = \{w \in s \mid \exists k : s[\phi]^g_k \neq \emptyset\}; \\
s[\partial\phi]^g_h = t & \text{iff} & s = t \ \& \ s = s[\phi]^g_h.
\end{array}
$$

where $g[x_n]h$ iff $h(x_n) \in h(n)$ & $\forall v \in (N \cup \mathcal{V}_N) : v \neq x_n \Rightarrow g(v) = h(v)$ (note that for all CC-indices n: $h(n) = g(n)$).[37]

[36] Note that assignments are total functions here. You can choose to have partial assignments, if you wish, but you don't have to.

[37] Note that variables can be reset. In the present formalization, this feature does not cause any 'downdate' problem (see footnote 11).

I will write $\exists x_n\phi$ to denote $\exists x_n \wedge \phi$, and, as usual, $\forall x_n\phi$ for $\neg\exists x_n\neg\phi$; $\Box\phi$ for $\neg\Diamond\neg\phi$; and $\phi \to \psi$ for $\neg(\phi \wedge \neg\psi)$. This implies the following semantics for $\to$:

$$s[\phi \to \psi]^g_h = t \quad \text{iff} \quad g = h \;\&\; t = \{w \in s \mid \forall k : w \in s[\phi]^g_k \;\Rightarrow\; \exists r : w \in (s[\phi]^g_k)[\psi]^k_r\}$$

It is easy to see that the following version of the update property holds now:

3.6.7. PROPOSITION. [Update property] $s[\phi]^g_h \subseteq s$

The evaluation process takes place on two levels: on the level of the states, where it is eliminative (as in the original formulation of Update Semantics of Veltman); and on the level of the assignments, where it proceeds in a point-wise relational fashion (as in the original formulation of Dynamic Predicate Logic of Groenendijk and Stokhof). This separation of levels is illuminating, because it shows the important difference between world and discourse information[38] which is assumed by the *CC* style of quantification. While factual information is functional, discourse information can yield more than one output.

I now define the notions of support and coherence.

3.6.8. DEFINITION. [CC-Support] $s \approx\!\!|_g\, \phi$ iff $\exists h : s = s[\phi]^g_h$

3.6.9. DEFINITION. [CC-Coherence] $\approx\!\!|_{CC}\, \phi$ iff $\exists M, g, \exists s \in S_M$: $s \neq \emptyset \;\&\; s \approx\!\!|_g\, \phi$

It is easy to see that the notion of CC-support is 'compositional'. If a state supports $\phi \wedge \psi$, then ψ must be supported in some intermediate state resulting from a successful update with ϕ (the proof is straightforward and relies on the update property):

3.6.10. PROPOSITION. [Support of Conjuncts]

$$s \approx\!\!|_g\, \phi \wedge \psi \;\Rightarrow\; s \approx\!\!|_g\, \phi \;\;\&\;\; \exists h : s[\phi]^g_h \approx\!\!|_h\, \psi$$

In the present semantics, we manage to match Dekker's predictions on when a speaker is licensed to utter a certain proposition.[39] You are not licensed to utter $\phi \wedge \psi$, if you are not licensed to utter ψ after ϕ. As we saw, this property of CC-support allows us to solve the underspecification 2 difficulties. Indeed, the wrong suspect sentence is not coherent in CC:

(165) $\not\approx\!\!|_{CC}\, \exists x_n \Diamond P x_n \wedge \neg P x_n$

The following three sentences are instead coherent and this shows that we avoid the underspecification 1, overspecification 1 and 2 respectively:

[38] On this issue see van Eijck and Cepparello (1994).
[39] See Dekker (1997) and Dekker (2000b).

(166) $\approx\!\!\!\mid_{CC} \exists x_n \Diamond P(x_n) \wedge \exists x_n \neg \Diamond P(x_n)$

(167) $\approx\!\!\!\mid_{CC} \exists x_n P(x_n) \wedge \forall y_m \Diamond (x_n = y_m)$

(168) $\approx\!\!\!\mid_{CC} \forall x_n \forall x_m (\Diamond x_n = x_m)$

In what follows, I state the relation between the old and the new formulation of dynamic semantics under conceptual covers. By the old formulation I mean the MDPL style formulation with the logical notions restricted to CC-accessible states.

Comparison

The old and the new formulations of dynamic semantics under cover do not define the same logical notions because the new formulation does slightly better than the old one in connection with the possibility of reusing variables. Let $\models\!\!\approx_{old}$ be $\models\!\!\approx_{CC2}$ (the chosen definition of coherence in the old system), and $\models\!\!\approx_{new}$ be $\models\!\!\approx_{CC}$ (the presently defined one). We then obtain the following:

(169) $\not\models\!\!\approx_{old} \exists x \phi \rightarrow \exists x \phi$

(170) $\models\!\!\approx_{new} \exists x \phi \rightarrow \exists x \phi$

However, if we disregard these sorts of variable clashes, the two versions of the semantics are equivalent. I will show this for the notion of coherence, which is the only general notion we have introduced so far. I will say that ϕ is *novel in* σ, if updating σ with ϕ does not involve any violation of the novelty condition; and ϕ is *novel*, if there is a σ such that ϕ is *novel* in σ. The sentence in (169) and (170) is an example of a non-novel sentence. I will also say that ϕ is *safe* in σ, if it is novel in σ and all its occurrences of free variables are defined in σ. The following proposition states that the old and the new formulation of the CC-semantics define the same notion of coherence, if we disregard cases of reuse of variables.

3.6.11. PROPOSITION. Let ϕ be novel. Then

$$\models\!\!\approx_{old} \phi \quad \text{iff} \quad \models\!\!\approx_{new} \phi$$

One direction of the proof hinges on the fact that given a new state s, a new assignment g and a novel sentence ϕ, we can construct an old state σ and perspective $\wp$, such that if $s \models\!\!\approx_g \phi$ (new), then $\sigma \models\!\!\approx_\wp \phi$ (old). For the other direction, we show that also given an old state σ connected under a perspective $\wp$, we can find a new state s and a new assignment g such that for all ϕ, if $\sigma \models\!\!\approx_\wp \phi$ (old), then $s \models\!\!\approx_g \phi$ (new). For the proof of proposition 3.6.11, see appendix A.3.

Truth and Entailment

I define now the notion of truth in a state s with respect to an assignment g.

3.6.12. DEFINITION. [CC-Truth] $s \models_g \phi$ iff $\forall w \in s$: $\exists h$: $w \in s[\phi]_h^g$

The following proposition states the relation between truth and support.

3.6.13. PROPOSITION. Let ϕ be a sentence, s a state and g an assignment.

$$s \models\!\approx_g \phi \;\;\Rightarrow\;\; s \models_g \phi$$

The converse does not hold. As an illustration, consider again the wrong suspect example:

(171) $\exists x_n \Diamond P x_n \wedge \neg P x_n$

This sentence is true in a state $s = \{w_1, w_2\}$ that encodes the information that either individual a (in w_2) or individual b (in w_1) is P, but, as we saw, it is not supported in it (see section 3.5).

(172) $s \models_g \exists x_n \Diamond P x_n \wedge \neg P x_n \;\;\&\;\; s \not\models\!\approx_g \exists x_n \Diamond P x_n \wedge \neg P x_n$

From this example, it is clear that the notion of truth is not a suitable notion for the characterization of licensing. However, truth is the right notion for the definition of entailment.

3.6.14. DEFINITION. [CC-Entailment]

$$\phi_1, ..., \phi_n \models_{CC} \psi \;\text{ iff }\; \forall M, s, g, h : s[\phi_1 \wedge ... \wedge \phi_n]_h^g \models_h \psi$$

Under this definition the following proposition is easily proved:[40]

3.6.15. PROPOSITION. [Deduction Theorem]

$$\phi, ..., \phi_n \models_{CC} \psi \;\text{ iff }\; \phi_1, ..., \phi_{n-1} \models_{CC} \phi_n \rightarrow \psi$$

Note that if we had defined entailment in terms of support rather than truth, we would have lost the deduction theorem. In the next subsection we say something more about the relation between entailment, truth and support.

Here are some typical examples of valid dynamic entailments:

(173) $\exists x_n P(x_n) \models_{CC} P(x_n)$

(174) $\exists x_n P(x_n) \rightarrow Q(x_n), P(z_n) \models_{CC} Q(z_n)$

[40]The proof is an easy exercise and follows directly from the definitions of entailment and implication.

The following schemes, which rule the interaction between quantifiers and modal operators, are also valid:

(175) $\models_{CC} \Diamond\exists x_n P(x_n) \rightarrow \exists x_n \Diamond P(x_n)$

(176) $\models_{CC} \exists x_n \Diamond P(x_n) \rightarrow \Diamond\exists x_n P(x_n)$

(177) $\models_{CC} \Diamond\forall x_n P(x_n) \rightarrow \forall x_n \Diamond P(x_n)$

Whereas the following is an example of a non-valid formula:

(178) $\not\models_{CC} \forall x_n \Diamond P(x_n) \rightarrow \Diamond\forall x_n P(x_n)$

The failure in (178) shows that the present semantics avoids the cardinality problems discussed above.

I conclude the section by briefly discussing the issue of the context (in)dependence of dynamic logic sentences.

Context (In)dependence

Quantification under conceptual covers formalizes the intuitive idea that quantifiers in natural language range *over* individuals *under* a perspective. Which perspective you adopt plays a role only in specific cases. In what follows, I briefly describe under which circumstances perspectives do and do not play a role.

First of all, consider the following two examples of non-valid entailments:

(179) $\exists x_n \Box\, P(x_n) \not\models_{CC} \exists x_m \Box\, P(x_m)$

(180) $\forall x_n \Diamond\, P(x_n) \not\models_{CC} \forall x_m \Diamond\, P(x_m)$

Quantification into holistic operators is sensitive to shifts of conceptualization, and this feature allows us to account, for instance, for the butler situations.

Let's see now what happens in the absence of holistic operators:

(181) $\exists x_n P(x_n) \models_{CC} \exists x_m P(x_m)$

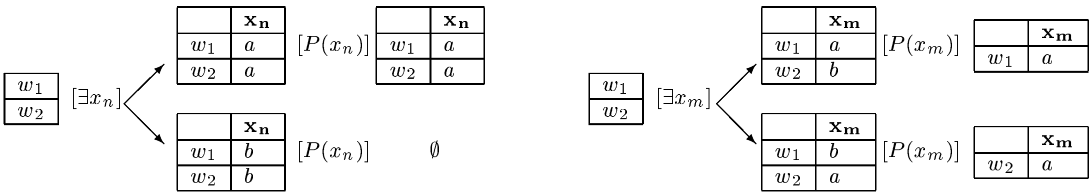
(182) $\forall x_n P(x_n) \models_{CC} \forall x_m P(x_m)$

If there are no holistic operators around, quantification is perspective independent with respect to truth and entailment, as we expected. On the other hand, note that in connection with the holistic notion of support, existential sentences, even if no holistic operator occurs in them, are sensitive to conceptual covers. Indeed, we can have a situation with one and the same state supporting one of the sentences in (183) and not supporting the other.

(183) a. $\exists x_n P(x_n)$

b. $\exists x_m P(x_m)$

The updates with the two sentences can have different effects in contexts where n and m are assigned different values. Let $s = \{w_1, w_2\}$ be such that individual a is the only P in w_1 and in w_2. Let g assign the rigid cover to n and another cover to m. The picture illustrates the two different results obtained by updating s with the two sentences with input assignment g.



Example (183a) is supported by s with respect to g and (183b) is not.

(184) $s \mid\!\approx_g \exists x_n P(x_n)$ & $s \not\mid\!\approx_g \exists x_m P(x_m)$

As an intuitive illustration, let $A = \{\text{Alfred, Bill}\}$ be the cover assigned to n, and $B = \{\text{the butler, the gardener}\}$ be the cover assigned to m. Furthermore, let P be the property 'having come to your office today'. Then s above can be taken to represent the belief state of a subject K finding himself in a by now well-known butler situation. There are two men, Alfred and Bill. One is the butler, the other is the gardener. K does not know who is who. K further believes that Alfred came to your office today. Consider now the following sentence which is true in s:

(185) Someone came to your office today.

The result in (184) then shows that the CC analysis predicts that K is licensed to utter (185) under cover A, but not under cover B. This prediction is intuitively correct. Indeed, if the addressee asks: 'Who?', then K can answer 'Alfred' under A, but he would have to admit 'I don't know' under B.

Thus conceptual covers matter for constructions containing quantification into some holistic operator (cf. examples (179) and (180)) or introducing some new discourse item (cf. example (184)), but note that this is the case only in situations of partial information. In a situation of total information, conceptualizations lose their bite. A state of maximal information is formalized by a set containing a single world $\{w\}$, but, with respect to a single world, all different conceptualizations collapse into one (see chapter 4, corollary 4.1.4). It follows that shifts of index do not matter in such a situation:

3.6.16. Fact. Let s be such that $|s| = 1$, and h be an assignment.

$$s \models\!\approx_h Q\,\vec{x}_{\vec{n}}\,\phi \quad \text{iff} \quad s \models\!\approx_h Q\,\vec{x}_{\vec{m}}\,\phi\,[\vec{x}_{\vec{n}}/\vec{x}_{\vec{m}}]$$

The presence of contrasting perspectives is a sign of ignorance. The one who knows everything knows how to map the possible conceptualizations onto each other, and, therefore, since he has a unique perspective over the domain, he can quantify over the individuals disregarding how these individuals are identified.

To conclude, quantification under conceptual covers is quantification over individuals under a certain perspective. Which perspective you choose only plays a role in certain circumstances, namely, in situation of partial information, for expressions introducing some discourse item and for sentences involving quantification into the scope of a holistic operator. These are typically the constructions in our formal language that are used to represent linguistic phenomena involving some notion of aboutness, such as *de re* attitude attributions, knowing-who constructions and specific uses of indefinite NPs. When we talk *about* individuals in situations of *partial* information, we do it under a conceptualization.

3.7 Conclusion

The combination of dynamic quantification with holistic notions is a dim affair, because it adds to the obscurity of quantification into modal contexts[41] problems typical of dynamic environments. In this chapter, I have tried to show that by bringing conceptual covers into the picture, we don't add obscurity to obscurity, but we shed some light on these difficult issues.

[41] See the concluding remarks in Quine (1956).

Chapter 4

Formal and Pragmatic Aspects of Conceptual Covers

In front of you lie two cards. One is the Ace of Spades, the other is the Ace of Hearts. Their faces are turned over. You do not know which one is which. There are two different ways of identifying the two cards in our domain, namely by their position on the table (the card on the left, the card on the right) and by their suit (the Ace of Spades, the Ace of Hearts). These identification methods are typical examples of what I have called conceptual covers in the previous chapters. In this chapter, I investigate the notion of a conceptual cover in more detail. In section 4.1, I restate the definition and discuss some general properties. I then investigate which view of trans-world identification is formalized by the notion of a conceptual cover and compare it with other views that have been proposed in the literature (section 4.2). Finally I will attempt a first analysis of the pragmatic procedures of conceptual cover selection (section 4.3).

4.1 Conceptual Covers

Given a set of worlds W and a set of individuals D, an *individual concept* is a total function from W to D. I call the *instantiation* of c in w the value $c(w) = d$ that a concept c assigns to a world w. A conceptual cover is a set of individual concepts which satisfies the following condition: in each world, each individual constitutes the instantiation of one and only one concept.

4.1.1. DEFINITION. Let W be a set of possible worlds and D a set of individuals. A *Conceptual Cover* CC over (W, D) is a set of individual concepts such that:

(i) $\forall w \in W : \forall d \in D : \exists c \in CC : c(w) = d$;

(ii) $\forall w \in W : \forall c, c' \in CC : c(w) = c'(w) \Rightarrow c = c'$.

The existential condition says that in a cover, each individual is identified by means of at least one concept in each world, that is, in each world, each individual should be accessible in some way. The uniqueness condition says that in no world, an individual is the instantiation of more than one concept, that is, in no world an individual is counted twice. This means that in a cover, either c_1 and c_2 are identical (their values coincide everywhere) or they are separated (their values never coincide). In a conceptual cover, each individual in the universe of discourse is identified in a determinate way, and different conceptual covers constitute different ways of conceiving of one and the same domain.

In the following two sections we prove two results, which have already been discussed in a number of applications in the previous chapters. The first proposition says that there is a one-to-one correspondence between a conceptual cover and the domain of individuals; the second proposition says that the number of covers definable over a set of worlds and a set of individuals is relative to the cardinality of these two sets. Technically these two result are not so impressive, but their intuitive consequences are interesting.

one to one

There is a one-to-one correspondence between a conceptual cover and the domain of individuals.

4.1.2. PROPOSITION. Let W be a set of worlds and D a set of individuals. Let CC be a conceptual cover based on (W, D). Then $|CC| = |D|$.

proof: Let f be a function from CC to D such that for some world w, the following holds: $f(c) = c(w)$. It is enough to show that f is a bijection. That f is injective and surjective follow directly from the uniqueness and the existential condition respectively:

(i) f is injective. $\forall c_1, c_2 \in CC : \; f(c_1) = f(c_2) \;\; \Rightarrow c_1(w) = c_2(w)$ (by construction) $\Rightarrow c_1 = c_2$ (by uniqueness condition).

(ii) f is surjective. $\forall d \in D : \exists c \in CC : c(w) = d$ (by existence condition) $\Rightarrow \forall d \in D : \exists c \in CC : f(c) = d$ (by construction). □

In a conceptual cover, each individual is identified by one and only one concept. Different conceptual covers constitute different ways of identifying individuals of the same sort. Irrespective of which perspective you assume, the number of individuals does not change.

A first consequence of proposition 4.1.2 is that in defining our models, we can drop the one domain assumption, as long as the domains associated with the individual worlds have the same cardinality.

Given this one to one correspondence between the universe of a model M and any conceptual cover based on M, any discussion about their relative priority becomes vacuous. In chapter 2, section 2.4.4, we have proved the related result that given a model M and a conceptual cover CC based on M, we can always find another model M' which satisfies the same sentences of M in which the identification method formalized by CC is the rigid one. The contrast between rigid and non-rigid conceptual covers can be used to represent the contrast between demonstrative versus descriptive identification. In this perspective, the one-to-one result can be interpreted as saying that our logic remains neutral with respect to the debate between empiricists and rationalists about the relative priority of these two identification methods. In the construction of the objects of our experience, the direct ostensive moment can be prior to the descriptive conceptual moment or vice versa, or the two moments can be taken to presuppose each other. We are free to embrace any of these philosophical positions. Logically speaking, it does not matter.

relativity

The number of conceptual covers definable over a set of worlds and a set of individuals crucially depends on the cardinality of the two sets. There are $(|D|!)^{|W|-1}$ conceptual covers over (W, D).[1]

4.1.3. Proposition. Let $C(W, D) = \{CC \mid CC \text{ is a cover based on } (W, D)\}$. Then $|C(W, D)| = (|D|!)^{|W|-1}$.

proof: the proof is by induction on the number of worlds $|W| = n$.

> $(n = 1)$ Let $|W| = 1$. Suppose $W = \{w\}$. Then there is a unique cover based on (W, D), namely $CC = \{\langle w, d\rangle \mid d \in D\}$. Thus $|C(W, D)| = 1 = (|D|!)^0 = (|D|!)^{|W|-1}$.
>
> $(n \Rightarrow n + 1)$ Let $|W| = n + 1$. We have to prove that $|C(W, D)| = (|D|!)^n$. Suppose $W = \{w_1, ..., w_n, w_{n+1}\}$. Consider $W^* = \{w_1, ..., w_n\}$ i.e. $W^* = W - \{w_{n+1}\}$. By induction hypothesis $|C(W^*, D)| = (|D|!)^{n-1}$. We have to check how many more conceptual covers we can have if we add w_{n+1} to W^*. Let $|D| = m$ and $CC^* = \{c_1, ...c_m\} \in C(W^*, D)$. Then there are $m!$ sequences of individuals $d_1, \ldots, d_m$, such that the conceptual cover CC, defined by
>
> $$CC = \{c_1 \cup \{\langle w_{n+1}, d_1\langle\}, \ldots, c_m \cup \{\rangle w_{n+1}, d_m\rangle\}\}$$
>
> is an extension of CC^* to $C(W, D)$. Since this holds for each CC^* in $C(W^*, D)$, and since $|C(W^*, D)| = (|D|!)^{n-1}$, then there are $(|D|!)^{n-1} \times |D|! = (|D|!)^n = (|D|!)^{|W|-1}$ in $C(W, D)$. □

[1] I am indebted to Rosella Gennari and Paul Dekker for the proof of this result.

The number of possible covers grows exponentially to the number of possibilities taken into consideration. The smaller a set of worlds, the less conceptual covers are available. In an epistemic perspective, this means that the number of non equivalent ways of identifying the objects in one domain is relative to the number of possibilities compatible with the relevant information state. The more informed a state, the less identification methods are available. A corollary of proposition 4.1.3 is that with respect to a singleton set of worlds there is only one possible conceptual cover.

4.1.4. Corollary. If $|W| = 1$, then $|C(W, D)| = 1$ for all D.

A singleton set of worlds represents the belief state of a subject who has a complete picture of what is the case, a subject without doubts about individuals and their properties. With respect to such a state, it is irrelevant which conceptual perspective you assume, because all different ways of specifying objects collapse into one. The availability of a number of non-equivalent methods of identification is a sign of lack of information.

The relativity result shows that identification methods are not given once and for all, but depend on a set of possible worlds. Indeed, as we have already seen numerous times in the previous chapters, sets of concepts which constitute proper conceptualizations with respect to a set of possibilities may cease to do so with respect to a larger set. This is due to the fact that if we add a world, two concepts, which had distinct values with respect to all old worlds, can overlap in the new world, and so can no longer be part of one and the same conceptualization.

As an illustration, consider again the card situation above. In front of you lie two cards turned over. One is the Ace of Hearts, the other is the Ace of Spades. You don't know which is which. Furthermore, the card on the left is marked and you know it. We can model your information state σ as follows:

$$\begin{array}{lll} w_1 & \mapsto & \heartsuit^{\bullet} \quad \spadesuit \\ w_2 & \mapsto & \spadesuit^{\bullet} \quad \heartsuit \end{array}$$

There are two conceptual covers based on σ:

$\mathrm{A} = \{\lambda w[\mathit{left}]_w, \lambda w[\mathit{right}]_w\}$

$\mathrm{B} = \{\lambda w[\mathit{Spades}]_w, \lambda w[\mathit{Hearts}]_w\}$

A identifies the cards by their position on the table and B identifies the cards by their suit. Which of these two methods of identification is operative, can influence evaluation. Consider, for instance, the following three examples:

(186) Which card is the Ace of Hearts?

$?x_n\ (x_n = \heartsuit)$

(187) Any card might be the Ace of Hearts.

$\forall x_n \Diamond (x_n = \heartsuit)$

(188) One of the cards is marked.

$\exists x_n ((x_n)^\bullet)$

If we assume cover A as value for n, then the question (186) partitions σ in two blocks (see chapter 1), the sentence (187) is consistent (see chapter 3), and the existential sentence (188) is supported by the state, so you are licensed to utter it (see again chapter 3). On the other hand, under a perspective which assigns the cover B to n, (186) is vacuous, (187) is inconsistent, and (188) is not supported by σ.

Suppose now you learn that the card on the right is the Ace of Hearts. Your information states will now look like the following, call it τ:

$$w_2 \quad \mapsto \quad \spadesuit^\bullet \quad \heartsuit$$

In such a situation, the two identification methods by suit and by position collapse into one. A and B restricted to w_2 are one and the same set, namely the set consisting of the following two functions from $\{w_2\}$ to $\{\spadesuit, \heartsuit\}$: $(w_2 \rightarrow \spadesuit)$ and $(w_2 \rightarrow \heartsuit)$. With respect to the new state τ, choosing different conceptual perspectives does not change the update effect of the three sentences above. Under any perspective, (186) is vacuous, (187) is inconsistent and (188) is supported in τ. The first two sentences are not acceptable, because both are intended to express gaps of information, and in a state of maximal information such as τ there are no such gaps. The existential sentence (188) is instead supported irrespective of which value is assigned to n. While in the previous situation, if you had chosen the wrong method of identification, you could have failed to be licensed to utter (188), – the sentence was not supported in σ under the identification by suit –, here the correctness of your utterance is not relative to the ways of specifying the cards in the domain. In a state of maximal information, you can quantify directly over the individuals disregarding the ways in which these are identified.

Suppose now Ralph enters the room and he knows that the card on the left is marked, but he believes that the Ace of Hearts is on the left, instead of the Ace of Spades. The following diagram visualizes Ralph's information state:

$$w_1 \quad \mapsto \quad \heartsuit^\bullet \quad \spadesuit$$

Suppose now you utter the following *de re* sentence:

(189) Ralph believes the Ace of Spades to be marked.

$\exists x_n (x_n = \spadesuit \wedge \Box_r (x_n)^\bullet)$

As soon as another subject enters the picture, conceptual covers are no longer irrelevant. The evaluation of (189) in w_2, which from your perspective stands for the actual world, involves taking into consideration also world w_1, which is the only possibility in Ralph's belief state (see chapter 2). With respect to $\{w_1, w_2\}$, A and B are not one and the same set of concepts. The two methods of identification by position and by suit are no longer interchangeable. Indeed, if n is assigned A, (189) is true; if n is assigned B, (189) is false.

To conclude, in a one agent situation, the presence of different contrasting perspectives over one domain is a sign of ignorance. The process of information increase can be characterized as a process of identification of different conceptualizations. In a state of maximal information all different perspectives coincide and we can talk directly about the individuals in such a state, independently from the ways in which these are specified. However, as soon as another agent enters the picture, different ways of identifying the object in the domain may arise and conceptual covers become relevant again to relate the individuals figuring in the different agents' epistemic alternatives.

4.2 Trans-world Identification

Methods of trans-world identification are 'ways of understanding questions as to whether an individual figuring in one possible world is or is not identical with an individual figuring in another possible world.'[2] In this section, I investigate which view of trans-world identification is formalized by the notion of a conceptual cover and I compare it with other views that have been proposed in the literature.

The most general way to define a method of cross-world identification is as a relation between world-individual pairs.

4.2.1. Definition. Let W be a set of possible worlds and D a set of individuals. A *method of cross-identification* R for $\langle W, D\rangle$ is defined as follows:

$$R \subseteq (W \times D)^2$$

A method of cross-identification tells you which individual is which across the boundaries of different possible worlds. We write $\langle w, d\rangle R\langle w'd'\rangle$ to indicate that d in w is identified with d' in w', or that d in w is the counterpart of d' in w'.

A typical example of a method of cross-identification is the following:

$$R_D = \{\langle w, d\rangle, \langle w', d'\rangle \mid d = d' \;\&\; w, w' \in W\}$$

R_D represents the view of trans-world identification presupposed by the G&S logic of question (see chapter 1), classical Modal Predicate Logic (see chapter 2)

[2]Hintikka (1969), p. 33.

and Slicing (see chapter 3). Two individuals are cross-identified iff they are one and the same individual. I call this view, the *rigid* method of cross-identification.

Any set of individual concepts determines a method of cross-identification. Given a set $S \subseteq D^W$, we can define the corresponding R_S as follows:

$$\langle w, d\rangle R_S \langle w', d'\rangle \text{ iff } \exists c \in S : c(w) = d \;\&\; c(w') = d'$$

By such a construction, the set of all concepts $IC = D^W$ determines a cross-identification method which identifies all individuals in one world with all others in each other world.

$$R_{IC} = \{\langle w, d\rangle, \langle w', d'\rangle \mid w = w' \rightarrow d = d'\}$$

R_{IC} is the method of cross-identification assumed by the first contingent identity semantics CIA in chapter 2. Any two individuals figuring in two different worlds are cross-identified. In this form, it strikes us for its intuitive inadequacy.

Conceptual covers are sets of concepts in which no splitting or merging is allowed. Each individual in one world is identified with one and only one individual in each other world. Clearly, also a conceptual cover determines a method of cross-identification. The converse obviously does not hold, not all methods of cross-identification determine conceptual covers. In this section, I am interested in specifying in a precise way the class of cross-identification methods R which do correspond to a conceptual cover. It turns out that a method of cross-identification R has to satisfy the most stringent conditions in order to determine a conceptual cover, namely:

(i) R is an equivalence relation:

  (a) R is reflexive;
  (b) R is symmetric;
  (c) R is transitive.

(ii) Each individual has one and only counterpart in each world:

  (a) $\forall w, w', d : \exists d' : \langle w, d\rangle R \langle w', d'\rangle$;
  (b) $\forall w, w', d, d', d'' : \langle w, d\rangle R \langle w', d'\rangle \;\&\; \langle w, d\rangle R \langle w', d''\rangle \;\Rightarrow\; d' = d''$.

I will call a method of cross-identification *proper*, if it satisfies (i) and (ii). The rigid method R_D above is a typical example of a proper R. Cross-identification methods R_S determined by arbitrary sets of concepts, are symmetric, but need not be proper. R_{IC} is an example of a non-proper method of cross-identification. Although R_{IC} is reflexive, symmetric and satisfies (iia), it does not satisfy transitivity and condition (iib). The methods of cross-identification A and B below are examples of relations satisfying condition (i) and (iia), but not condition (iib), and of relations satisfying (i) and (iib), but not (iia), respectively.

$A = (W \times D)^2$

$B = \{\langle w, d\rangle, \langle w, d\rangle \mid d \in D \;\&\; w \in W\}$

In A, all individuals in all worlds are identified. In B, no two individuals from two different worlds are identified.

Finally, to see that conditions (i) and (ii) are independent consider the counterintuitive method C which satisfies (ii), but is not an equivalence relation. Let $|W| = |D| = 2$:

$C = \{\langle w, d\rangle, \langle w', d'\rangle \mid d \neq d'\}$

I will show now that a method of cross-identification R corresponds to a conceptual cover iff R is proper. For more detailed proofs see Appendix A.4.

4.2.2. Proposition. Let CC be a conceptual cover over (W, D). The method of cross-identification R_{CC} determined by CC is proper.

Recall that R_{CC} is defined so that $\langle w, d\rangle R_{CC} \langle w', d'\rangle$ iff $\exists c \in CC : c(w) = d$ and $c(w') = d'$. The existence condition on conceptual covers implies that reflexivity and condition (iia) are satisfied (together with the fact that we are dealing with total functions). The uniqueness condition on CCs implies transitivity and condition (iib). Symmetry follows by construction of R_{CC}.

Let's see now how we can define a conceptual cover from a proper method of cross-identification.

4.2.3. Definition. Let R be a cross-identification method over (W, D). The set of classes of pairs induced by R is the following set:

$$CP_R = \{[w, d]_R \mid w \in W \;\&\; d \in D\}$$

where $[w, d]_R = \{\langle w', d'\rangle \mid \langle w, d\rangle R \langle w', d'\rangle\}$.

4.2.4. Proposition. The set of classes of pairs induced by a proper cross-identification method R is a conceptual cover.

This result follows directly from the two lemmas below.

4.2.5. Lemma. Let R be a proper method of cross-identification over (W, D). Then

$$\forall \alpha \in CP_R : \forall w \in W : \exists! d \in D : \langle w, d\rangle \in \alpha$$

Lemma 4.2.5 states that if R is proper, then each element of CP_R uniquely determines a total function from W to D, that is an individual concept.

I write $\alpha(w)$ to denote the individual d such that $\langle w, d\rangle \in \alpha$.

4.2.6. LEMMA. Let R be a proper method of cross-identification over (W, D). Then

$$\forall w \in W : \ \forall d \in D : \ \exists! \alpha \in CP_R : \alpha(w) = d$$

The proof of this lemma follows directly from the fact that if R is an equivalence relation, then CP_R is a partition of the set of world-individual pairs. Lemma 4.2.6 states that the set of concepts CP_R satisfies the uniqueness and the existence conditions and therefore is a conceptual cover.

Each conceptual cover uniquely determines a proper method of cross-identification (proposition 4.2.2) and each proper method of cross-identification uniquely determines a conceptual cover (proposition 4.2.4).

The existence and uniqueness conditions, which characterize a conceptual cover, thus find further justification from this perspective. Condition (i) is a quite natural constraint on methods of identification. In each world, each individual should be identified to itself (reflexivity). If d in w is identified with d' in w', then d' is w must be identified with d in w (symmetry). Finally, if d in w is identified with d' in w' and d' in w' is identified with d'' in w'', then also d in w and d'' in w'' must be identified (transitivity).

Condition (ii) is also an intuitive constraint. On the one hand, it simply says that two individuals cannot become one. On the other, it requires that no individual can cease to exists once we move from one world to the other.

The rigid view of trans-world identification exemplified by R_D above, being proper, corresponds to a conceptual cover. However, many other methods of identification are proper as well. In the analysis I defend in this thesis, different proper methods are allowed to act as the operative ones in different occasions.

4.2.1 Alternative Views of Trans-world Identification

After having seen which view of trans-world identification is formalized by the notion of a conceptual cover, I will now review a number of alternative views that have been proposed in the literature. The comparison will also give us the occasion to discuss a series of future applications and loose ends of the present analysis.

Individuating Functions

The present thesis can be seen as a development of a simple insight that Jaakko Hintikka presented in two articles at the end of the 60s where he envisaged the availability of different methods of cross-identification on different occasions.[3] In these articles, however, Hintikka did not carry his own insight far enough. The empirical applications he discusses are not totally convincing and the formalization he assumes is quite unsatisfactory. Hintikka (1969) discusses the logic of

[3]See Hintikka (1967) and Hintikka (1969).

propositional attitudes. Hintikka recognizes that quantification in such a logic presupposes a method of cross-identification and formalizes such methods by sets F of (possibly partial) individuating functions f mapping worlds to individuals, satisfying the following non-splitting condition: $\forall w : \forall f, f' \in F : (f(w) = f'(w) \Rightarrow \forall w' : f(w') = f'(w'))$, which corresponds to my uniqueness condition. Although he implicitly proposes to relativize quantification to such methods of identification, Hintikka does not discuss empirical applications of such a relativization in connection with propositional attitude reports. For an example of such a concrete application, we have to go back to Hintikka (1967), where he discusses the semantics of perception reports. In this article, Hintikka isolates two distinct methods of cross-identification (perspectival or demonstrative cross-identification vs physical cross-identification) and introduces two kinds of quantifiers $\exists x$ and Ex, the former ranging over individuals identified by the first method, the latter ranging over individuals identified by the second method. This distinction is used to capture the contrast between direct and indirect perception reports of the following kind:

(190) d perceives Mr. Smith.

$\exists x(d$ perceives that Mr. Smith $= x)$

(191) d sees who the man in front of him is.

$Ex(d$ perceives that the man in front of $d = x)$

I will not discuss Hintikka's logic of perception that has been criticized (see Barwise and Perry (1983) for a more influential analysis of perception reports). Hintikka's insight that perceptive reports and propositional attitude reports can involve different methods of cross-identification is very impressive, but it should be generalized substantially. Hintikka's strategy to introduce a different quantifier for each method of cross-identification is not satisfactory. Obviously, far more than two methods are operative in our ordinary conversations, and so, such a strategy would lead to an intolerable complication of the syntax, while not explaining the context sensitivity involved in these cases.

Other authors have proposed a semantics involving quantification over sets of individual concepts, for instance Kraut (1983) and Zeevat (1995). I will just consider the former. Kraut (1983) discusses propositional attitude reports and proposes to analyze them in terms of quantification over a set of individuating functions but avoids Hintikka syntactic complications by making the domain of quantification explicitly context dependent. Kraut's approach constitutes a clear example of what I called a pragmatic analysis in chapter 2. Although they are close in spirit, the theory I defend in chapter 2 and Kraut's analysis depart on a series of points. Firstly, Kraut does not discuss any particular extra constraint on his sets of individuating functions (in particular no uniqueness condition is assumed) and, in addition, his analysis of *de dicto* attitude reports is quite different from mine. As announced in the title of the article, according to Kraut, there are no *de dicto* attitudes. Kraut proposes the '*de re*' representation in (193) for the traditional *de dicto* reading of a sentence like (192):

(192) Ralph believes that the tallest member of the club is a spy.

(193) $\exists x\ \Box_r(x$ is the tallest member of the club & x is a spy)

This means that on Kraut's account, traditional *de dicto* reports show the same context dependence as traditional *de re* reports and this seems to me at variance with our intuition. Consider a situation in which Ralph would assent to the sentence 'The tallest member of the club is a spy' while having no idea as to the identity of the tallest member of the club. According to the analysis in chapter 2, the *de dicto* reading of (192), which is represented in accordance with the tradition, by sentence (194), is true in such a situation irrespective of the operative method of identification.

(194) $\Box_r \exists x_n(x_n$ is the tallest member of the club & x_n is a spy)

On the other hand, on Kraut's account, the acceptability of (192) is relative to the conceptual perspective we assume. Indeed, in order for (193) to be true, the concept 'the tallest member of the club' must be part of the operative method of identification. Now, such a method is clearly not the prominent one in the described circumstances (Ralph has no idea who the tallest person in the club is, it is said), so a shift of identification method would be required in order to interpret such a simple and ordinary case. I don't think this is correct. Not all attitude reports are relative to a method of identification and the so-called *de dicto* readings are precisely those for which there is no such dependence.

The reason why Kraut assumes such an analysis of *de dicto* reports has to do with his main empirical application, namely intentional identity phenomena. In order to be able to represent Hob-Nob sentences like (195a) by a 'geachean' representation (195b), the existential quantifier must take wide scope also in traditional *de dicto* sentences:

(195) a. Hob thinks that a witch has blighted Bob's mare, and Nob believes that she [the same witch] killed Cob's sow.

b. $\exists x(\Box_h \phi \wedge \Box_n \psi)$

Kraut's solution to the intentional identity puzzle is not totally satisfactory though. For instance, if we can represent (195a) by (195b), then (196b) is also a possible representation for (196a), but then we lack an explanation of the unavailability of the anaphoric relation in the latter case:[4]

[4]Kraut might argue in favor of a pragmatic solution to this difficulty which involves the use of partial concepts. By assuming a method of identification containing a concept – the relevant

(196) a. Hob thinks that a witch has blighted Bob's mare. (?) She killed Cob's sow.

b. $\exists x(\Box_h \phi \wedge \psi)$

The difficulty I have indicated is just one of the many intricacies arising from the obscure area of intentional identity. Sets of individuating functions as well as conceptual covers can be used to determine which individuals corresponds to which in different agents' states. Therefore, Hob-Nob phenomena seem a natural application for a system using individual concepts. However, things are far more complex than they seem and, therefore, I must leave a proper treatment of these phenomena to another occasion.

Counterpart Relation

Lewis (1968) introduces the counterpart relation as a substitute for identity between things in different worlds. Lewis is discussing a metaphysical notion and not an epistemic or doxastic one. Metaphysical considerations lead him to assume disjoint domains, so that no individual is allowed to figure in two different worlds. Lewis' counterpart relation C is a relation on $\bigcup_{w \in W} D_w$, where dCd' means that d is the counterpart of d' in another world. Lewis assumes that C must satisfy the constraint that if d and d' are both in D_w, then dCd' iff $d = d'$. The counterpart relation is intuitively based on the relation of similarity between individuals figuring in different worlds which need be neither symmetric nor transitive.[5] On the other hand, each individual is similar to itself, so C is assumed to be reflexive. In our formulation, we can express Lewis counterpart relation by means of the following conditions on the method of cross-identification R:

1. $\forall w, d, d' : \langle w, d\rangle R \langle w, d'\rangle \Rightarrow d = d'$;
2. $\forall d{:}\exists w : \langle w, d\rangle R \langle w, d\rangle$;
3. $\forall w, w', d{:}\langle w, d\rangle R \langle w', d\rangle \Rightarrow w = w'$.

Condition 1 is Lewis' constraint mentioned above that in one and the same world, an individual can only have one counterpart, namely itself. Condition 2 corresponds to the constraint that C is reflexive. The last condition expresses the fact that no individual can occur in two different worlds (we can define the set of individuals D_w figuring exclusively in w as follows: $D_w = \{d \in D \mid \langle w, d\rangle R \langle w, d\rangle\}$). If we let R satisfy conditions 1, 2, and 3, we then obtain that dCd' iff $\exists w, w'$:

witch – which is defined in all Hob's and Nob's possibilities, but undefined in the actual world, we can account for the acceptability of (195a) and the unacceptability of (196a). However, the contrast between the two sentences seems to me to be more structural than a question of pragmatics. Opinions about this issue may diverge though.

[5] Lewis (1968), pp. 115-116.

$\langle w, d\rangle R \langle w', d'\rangle$. Notice that conditions 1 and 2 are satisfied also by conceptual covers (condition 2 is reflexivity and condition 1 follows from reflexivity and condition (iib)). Condition 3 is instead not needed if we assume an epistemic perspective. Conceptual covers further satisfy symmetry, transitivity and the condition that each individual has one and only one counterpart in each world, for which I have already argued for. Counterpart relations have also been assumed in connection to epistemic or doxastic phenomena by Lewis himself in (1983), but also in Stalnaker (1987) and van Rooy (1997). In the latter, counterpart theory is used to account for the double vision puzzles of *de re* belief attributions. Van Rooy proposes to analyze these constructions by means of quantification over sets of counterpart functions.[6] The analysis I defend in chapter 2 and van Rooy's semantics of belief reports have the same empirical coverage, at least as far as double vision cases are concerned.[7] On the other hand, from a formal point of view, the two analyses are quite different, and mine is more conservative in staying as close as possible to ordinary modal predicate logic. This feature has many advantages in terms of simplicity and tractability of the logic, e.g. the empirical predictions of the system are easily seen and its formal properties, notably the completeness result, can be proved by standard techniques. This is not always the case in van Rooy's system.

Individuation Schemes

In two short articles,[8] Jelle Gerbrandy presents, in a lucid way, the central idea I defend and investigate in this thesis: different methods of identification are operative in different contexts and evaluation of fragments of the discourse can vary relative to these methods. Gerbrandy also discusses the question of cross-identification in connection with epistemic phenomena, in particular knowing-who constructions, and proposes to relativize their interpretation to what he calls *individuation schemes*. These are methods of cross-identification R where R is an equivalence relation.

In chapter 1, section 1.4.3, I have already argued in favor of the adoption of

[6] Van Rooy's counterpart functions C map individual-world pairs (d, w) to individuals d' element of D_w. By adopting counterpart functions rather then relations he rules out the possibility of one individual having two counterparts in some world (see condition (iib) above). Van Rooy can still express double vision cases, which crucially involve such splitting, by taking sets of possibly 'overlapping' counterpart functions rather than single counterpart relations.

[7] I am not sure about the predictions of van Rooy's analysis in relation to the other problems I have discussed in chapter 2, e.g. the shortest spy problems, but I guess my analysis does slightly better in connection with these cases. Although van Rooy recognizes the context sensitivity of belief attributions, this is not reflected in his semantics where he seems to follow Kaplan and encodes information about which counterpart functions are suitable for quantification in the model (see his notion of counterpart functions by acquaintance C_{acq}). But, as I argued in chapter 2, such a strategy cannot solve the shortest spy problems without generating other difficulties (see the cases of Odette's lover and Susan's mother in chapter 2, section 2.3.2).

[8] Gerbrandy (1997) and Gerbrandy (2000).

more stringent constraints[9] for R in connection with questions and knowing-who constructions. My argument there had to do with the intuitive relation between knowing-who and knowing-how many constructions. If we relativize questions to possibly non-proper identification methods, we predict that one can know who is P, without knowing how many are P. The following sentence would be consistent, and this would be at variance with our intuition:

(197) I don't know how many people were late today, but I know who was late today.

It could be argued, however, that there are interpretations under which this sentence might be acceptable, namely, if we interpret the second conjunct as saying that you know which *kinds* of individuals were late today, rather than which *individuals*. For instance, you could perhaps consistently say:

(198) I don't know how many were late today, but I know who was late today, namely some linguists and some logicians.

A possible justification for dropping condition (ii), which is otherwise quite intuitive, is that by taking equivalent cross-identification methods R in which two or more individuals can become one, we can account for this individual-kind ambiguity. However, the individual-kind distinction seems to be of a different nature than the one formalized by different covers. While the latter involves different (though equivalently fine-grained) ways of identifying the entities in the domain, the former involves looking at one domain assuming different levels of granularity. Gerbrandy's individuation schemes might be used to formalize such differently fine-grained ways of conceiving one domain. Ginzburg (1995) accounts for examples like (198) by adopting a relative notion of answerhood.[10] By assuming that different contexts can select different individuation schemes as quantificational domains for the wh-phrases, we might be able to deal with the individual-kind distinction while maintaining an ordinary notion of answerhood. However, spelling out the details of such an analysis is not a trivial task. Furthermore, it is not totally clear to me whether a domain selection strategy is the correct one here, since there is a lot of vagueness playing a role, and therefore an analysis like that in Ginzburg (1995) might be more appropriate for these cases.

4.3 Towards a Pragmatic Analysis

The main idea of the analysis I defend in this thesis is that different methods of identification are available on different occasions and that evaluation of sentences

[9] Namely R should correspond to a conceptual cover, that is, it should be a proper method of cross-identification, i.e, it should also satisfy condition (ii) and not only be an equivalence relation.

[10] See chapter 1, section 1.5.3 for a discussion of Ginzburg's analysis.

can vary relative to these methods. I have proposed to characterize methods of identification by means of the notion of a conceptual cover and I have studied a number of linguistic constructions whose interpretation depends on the contextually selected conceptualization, notably questions and knowing-who constructions (chapter 1), belief attributions (chapter 2), epistemic modals, presupposition and specific indefinite NPs (chapter 3). Such constructions can express different meanings in different contexts. The question I will explore in the remaining part of this chapter is how the addressee may be able to select the intended identification method while interpreting these constructions. In order to shed some light on this complicated issue, I will use notions from Optimality Theory (OT) and Game Theory (GT). The analysis I propose in the following pages is still in its germinal phase, and needs further investigation. Nevertheless, it shows interesting aspects of the cover selection procedure, the most significant one being that shifts of cover never occur without justification. Furthermore, it illustrates that the use of OT and GT notions in the explanation of phenomena lying on the semantics-pragmatics interface is promising, although not totally unproblematic.

According to an OT analysis of interpretation (see de Hoop and Hendriks (1999)) the process of interpretation of natural language sentences is ruled by a number of ordered interpretation constraints. The addressee chooses from a set of possible meanings the ones which optimally satisfy these constraints. A simple machinery of this sort is sufficient to give a rational explanation of the cover selection procedures involved in many of the examples we have considered in the previous chapters, but not all. A number of potentially problematic cases of *de re* sentences cannot be explained by such an addressee-oriented analysis. I will suggest that a proper treatment of such examples requires a bi-dimensional interpretation theory (see Blutner (1999) and Blutner and Jäger (1999)), in which also the speaker's perspective is taken into consideration. In such a theory, the optimal solution is searched along two dimensions, the one of the addressee and the one of the speaker whose choice of uttering this or that sentence is influenced firstly by general principles of generation, secondly by the principle of cooperation, and finally by her particular interests and goals. I will follow Dekker and van Rooy (1999) and recast bi-dimensional OT interpretation processes in terms of 'interpretation games'. Game Theory turns out to be a promising framework for describing the interplay of general linguistic constraints and particular goals in the search for an optimal interpretation.

4.3.1 OT Interpretation Theory

In Optimality Theory (see Prince and Smolensky (1997)) conflicts between constraints are arbitrated by ranking one constraint over the other. OT has been applied in phonology, where it constitutes the dominant theoretical paradigm, in syntax, and, recently, also in semantics and at the semantics-pragmatics interface (see de Hoop and de Swart (1999), de Hoop and Hendriks (1999), Blutner (1999), Blutner and Jäger (1999), Zeevat (1999a), Zeevat (1999c) and Dekker and van Rooy (1999).)

An OT interpretation theory is based on a set of constraints ordered according to their relative strength, which help us in deciding between different readings allowed by the generative part of the grammar. The addressee has a set of alternative contents for a specific expression at her disposal. The best interpretations are those elements of the set which do better on the interpretation constraints than all other alternative candidates, where candidates that have arbitrary many violations of lower ranked constraints do better than candidates that have also one violation of a higher ranked constraint.

The phenomena that we have considered in the previous chapters provide evidence in favor of competition and ranking of interpretation constraints. In the following pages, I will briefly discuss a number of constraints which seem to play a role in the process of cover selection. I will then present examples of conflicts between these constraints and I will discuss which ranking is suggested by these conflicts. In the discussion, I follow Blutner (1999) who adopts updates in an OT setting. Meanings are identified with information change potentials (see chapter 3). Information states are formalized as sets of possibilities. Assertions modify these states in various ways by eliminating possibilities or extending them. Questions are taken to partition the states in alternative blocks.[11] As in the previous chapters, sensitivity to methods of identification is expressed by means of CC-indices $n, m, \ldots$. I will write $UP^{\wp}_{\sigma}(\alpha[\vec{x}_{\vec{n}}])$ to denote the set of the potential outcomes of updating a state σ with an expression $\alpha[\vec{x}_{\vec{n}}]$ under perspective $\wp$. As in chapters 1 and 3, a conceptual perspective $\wp$ is a function assigning conceptual covers to the CC-indices $\vec{n}$ occurring in α.

Interpretation Constraints

The first interpretation constraint I will discuss is the principle of ANCHOR discussed in Zeevat (1999a).

ANCHOR says that interpretation should be anchored to the context. This principle governs the interpretation of expressions which are assigned a value either by deixis or by anaphora resolution, and hence should find a proper antecedent in the context. Examples of such expressions are pronouns, tenses and CC-indices. Normally antecedents for such expressions are made salient either by explicit mention in the preceding discourse (anaphora resolution) or by the actual presence of the relevant referents in the utterance situation (deixis). CC-indices are no exception to this. Consider again the Spiderman, Ortcutt and butler examples that have been presented in chapter 1, 2 and 3 respectively. Recall the relevant situations:

[11] See Groenendijk (1998) and Groenendijk (1999) for a dynamic treatment of interrogative sentences.

Spiderman Someone killed Spiderman. You have just discovered that John Smith is the culprit. So you can say (a). Now John Smith is attending a (masked) ball. You go to arrest him there, but you don't know what he looks like. So you say (b).

(199) a. (John Smith did it. So) I know who$_n$ killed Spiderman.

b. (This person might be the culprit. That person might be the culprit. So) I don't know who$_m$ killed Spiderman.

Ortcutt You can tell each half of the Ortcutt story separately. In one half Ralph sees Ortcutt wearing the brown hat. In the other he sees him on the beach. From the first story you can reason as in (a). From the second story you can reason as in (b).

(200) a. Ralph believes that the man with the brown hat is a spy.
The man with the brown hat is Ortcutt.
So Ralph believes of Ortcutt that he$_n$ is a spy.

b. Ralph believes that the man seen on the beach is not a spy.
The man seen on the beach is Ortcutt.
So Ralph does *not* believe of Ortcutt that he$_m$ is a spy.

the butler Suppose a butler and a gardener are sitting in a room. One is called Alfred and the other Bill. We don't know who is who. The butler is the culprit.

(201) a. The butler did it. So it is not true that anybody$_n$ in the room might be innocent.

b. Alfred might be innocent. Bill might be innocent. So anybody$_m$ in the room might be innocent.

In the (b) case of the Spiderman example, the value for the index m is suggested by the concepts given by the visual images of the masked faces, which become salient by entering the perceptual field of the participants in the conversation. In all other cases, one conceptualization or the other is suggested as value for the relevant index by the previous discourse which explicitly mentions one or the other concept. In all examples, the context supplies as antecedents for the CC-indices single isolated concepts rather than the conceptual cover themselves. In contrast with the case of pronouns, it seems that the context can contribute to determine the value of a CC-index n by merely suggesting conditions for this value rather than by supplying the value itself. These conditions have normally the following form:

$$\{c_1, ..., c_m\} \subseteq n$$

where $c_1, ..., c_m$ are concepts salient in the context. Such conditions can clearly fail to uniquely determine a specific conceptual cover. In these cases, the content expressed by a sentence $\phi[x_n]$ can fail to be uniquely determined. By the specification of a set of possible conceptualizations, a set of possible contents can be selected, rather than one. This indeterminacy, though, does not necessarily lead to failure of communication. The condition for avoiding such a failure is that all these contents behave uniformly with respect to the relevant background state. The point of an assertion is to reduce the background state in a certain determinate way. If the sentence is associated with a number of contents that affect the input state in different ways, it would be unclear whether a possibility should be eliminated or included in the resulting state. On the other hand, if all the contents involve the same action on the background state, no ambiguity arise (see Stalnaker (1978)). Therefore, we can assume that a condition restricting the possible values of a CC-index n that uniquely determines the effect of the utterance of a certain sentence $\phi[x_n]$ on the input state can act as a suitable antecedent for n. More formally, if for all perspectives $\wp$, $\wp'$ the following holds: $\{c_1, ..., c_m\} \subseteq \wp(n)$ & $\{c_1, ..., c_m\} \subseteq \wp'(n) \;\Rightarrow\; UP^{\wp}_{\sigma}(\phi(n)) = UP^{\wp'}_{\sigma}(\phi(n))$, then the sequence $c_1, ..., c_m$ can act as a suitable antecedent for n. In a context in which such a sequence of concepts is salient ANCHOR can be satisfied. As an illustration, consider the final *de re* sentence in the (a) case of the Ortcutt example:

(202) Ralph believes of Ortcutt that he$_n$ is a spy.

$$\exists x_n(x_n = o \wedge \Box S(x_n))$$

Let σ be any state resulting from an update with the two preceding sentences in (200a). The condition in (203) is sufficient for a felicitous interpretation of (202) in such a situation, because, in σ, the update brought about by (202) will be the same under any cover which satisfies (203):

(203) $\lambda w[\text{the man with the brown hat}]_w \in n$.

A condition like the following might instead cause indeterminacy:

(204) $\lambda w[\text{the shortest spy}]_w \in n$.

Since it may not supported by the background state σ that Ortcutt is the shortest spy, it would not be clear under which perspective to identify Ortcutt in Ralph's belief's state, so (202) can turn out both supported and rejected in σ.

Now, ANCHOR says that interpretation should be anchored. All anaphoric expressions should find a proper antecedent in the context and we have just discussed the peculiar modalities of the anchoring of CC-indices. Still assuming that the absence of a suitable salient antecedent for a CC-index leads to communication breakdown is quite unrealistic. In real life communication, people deal with these cases by accommodating one or the other (condition on) conceptualization.

So, in such situations we should allow accommodation. However, accommodation should be disallowed in case a proper antecedent is already available in the context. OT interpretation theory can capture the latter intuition by assuming a principle which prohibits the addressee the addition of new material to the context, Zeevat (1999a) calls such a principle ***ACCOMMODATION** (see also Blutner (1999)). But, OT can also account for the fact that accommodation is allowed in certain circumstances by positing that ANCHOR can overrule *ACCOMMODATION (see again Zeevat (1999a)). If no antecedent is available we choose to accommodate in order to satisfy ANCHOR which ranks higher than *ACCOMMODATION. If an antecedent is already present in the context, ANCHOR is satisfied. Consequently *ACCOMMODATION is the critical constraint and its violation become crucial. So, we prefer readings which do not involve accommodation.

ANCHOR in interaction with *ACCOMMODATION explains the contrast between the (a) and the (b) cases in the Spiderman, Ortcutt and butler examples. I will just consider the latter case:

(205) a. (The butler did it. So) it is not true that anybody$_n$ in the room might be innocent.

$\neg\forall x_n \Diamond I(x_n)$

b. (Alfred might be innocent. Bill might be innocent. So) anybody$_m$ in the room might be innocent.

$\forall x_m \Diamond I(x_m)$

Consider example (205b). The CC-index m should intuitively be assigned the cover $A = \{$Alfred, Bill$\}$ rather than the cover $B = \{$the butler, the gardener$\}$. By assuming ANCHOR and *ACCOMMODATION we explain this preference. A is salient in the context, whereas B is not. Hence, an assignment of m to B would violate *ACCOMMODATION, whereas an assignment of A to m would not involve any violation. By the same kind of reasoning, n is assigned cover B, and not A in (205a).

Another clear example in which ANCHOR plays a crucial role is the following case discussed in chapter 1.

Priscilla Consider sentence (206) uttered by Priscilla in the two situations α and β:

α: In your living room.

β: At a party with many African leaders.

(206) Who is the president of Mali?

a. Konare is the president of Mali.

b. He [pointing at Konare himself] is the president of Mali.

On the one hand, an interpretation of the question under a demonstrative cover containing Konare himself is clearly not anchored in context α and accommodating in this case would mean flying to Africa and kidnapping Konare. On the other hand, in context β, (206) can be interpreted under the demonstrative cover without violations. Indeed, (b) is a good answer to (206) in β, but not in α.

The next principle I will discuss is a constraint which specifically governs cover selection procedures.

***SHIFT** expresses a general preference for interpretations which do not involve shift of conceptualizations. As an illustration of the role of *SHIFT, consider the following variation on the workshop example discussed in chapter 1:

the workshop You are attending a workshop. In front of you lies the list of names of all participants, around you are sitting the participants in flesh and blood. You don't know who is who. Consider now the following question:

(207) Who$_n$ has taken whom$_m$ to the party yesterday night?

$?x_n y_m R(x_n, y_m)$

Although two conceptualizations are salient in the described situation, namely naming and the ostensive cover, there is a clear preference for a uniform interpretation for n and m. Indeed, in such a situation, where you do not know which person is called what, replies like (208) or (209) are intuitively more acceptable answers to (207), than a reply like (210):

(208) Dylan Dog has taken Nathan Never. Ken Parker has taken Dylan Dog. ...

(209) This man has taken that man. The man in the first row has taken that guy over there. ...

(210) This man has taken Nathan Never. That man has taken Ken Parker. ...

Before turning to cases in which *SHIFT is overruled, I will discuss the principle of STRENGTH.

STRENGTH is a constraint which specifically governs the process of cover selection and expresses a preference for the selection of stronger covers.[12] Intuitively, if A is a stronger cover than B, then A represents the objects of our experience better than B. Before trying to give a formal characterization of the notion of a stronger cover, I will discuss a somewhat artificial, but clarifying example, of which the interpretation seems to be ruled by such a constraint.

Consider the following situation. Suppose you have two neighbors, the upstairs neighbor who is called John Smith and who is the major of your town and the downstairs neighbor who is called Bill White and who is the local mafia boss. You also have two roommates K and J. Suppose now that (a) K knows that John Smith is the major, and Bill White is the local mafia boss, but he has no idea that they are living in your own apartment block; (b) J knows that the major is living upstairs and the mafia boss downstairs but he has no idea what they are called.

Suppose now that John Smith has called your apartment for one or other reason. K and J come home at different times and ask you separately:

(211) Who$_n$ called?

The intuition is that not the same cover can be accommodated as value for n on the two occasions. For instance, you could accommodate naming as value for n in case K asked the question, but not in the other case. Indeed,

(212) K: Who$_n$ called?

a. John Smith.

b. (?) The upstairs neighbor.

c. The major.

(213) J: Who$_n$ called?

a. (?) John Smith.

b. The upstairs neighbor.

c. The major.

Let A be naming, B the cover containing the concepts 'the upstairs neighbor' and 'the downstairs neighbor', and C the third relevant cover containing the concepts 'the major' and 'the local mafia boss'. Now I will say that A and C are stronger than B in the first case, whereas B and C are stronger than A in the second

[12] This principle is different from Zeevat's principle of strength which expresses preference for informationally stronger readings. As will become clear soon, if A is stronger than B, an interpretation under cover A does not have to be informationally stronger than an interpretation under cover B.

case. This ranking is predicted by the definition of relative strength of a cover which I propose below according to which the strength of a cover depends on the number of alternative relevant covers that have been identified with it in the specific information states of the relevant agents. By assuming a constraint like STRENGTH we can account for our intuitions with respect to this example.

The general idea of the definition of relative strength is that 'clusters' of covers are stronger than isolated ones. The notion of a strong cover is highly context dependent. There are no absolute strong covers. A cover is stronger than another with respect to the set of covers that we take into consideration, and relative to a(n) (set of) information state(s).

I first define the auxiliary notion of a cover restricted to a set of worlds σ, and the notion of an equivalence class of covers in a set C with respect to σ.

4.3.1. Definition. Let CC be a cover based on (W, D), and $\sigma \subseteq W$. The restriction of CC to σ, CC_σ is defined as follows:

$$CC(\sigma) = \{c \in D^\sigma \mid \exists c' \in CC : \forall w \in \sigma : c(w) = c'(w)\}$$

4.3.2. Definition. Let C be a set of covers based on (W, D), $CC \in C$, and $\sigma \subseteq W$.

$$[CC]_{\sigma,C} = \{CC' \in C \mid CC(\sigma) = CC'(\sigma)\}$$

I define now the relation being at least as strong as, $\gg_{\sigma,C}$ on a set of covers C with respect to a set of worlds σ.

4.3.3. Definition. [Relative Strength] Let CC_1 and CC_2 be two covers in C based on (W, D), and $\sigma \subseteq W$.

$$CC_1 \gg_{\sigma,C} CC_2 \text{ iff } |[CC_1]_{\sigma,C}| \geq |[CC_2]_{\sigma,C}|$$

A cover which is identified with n relevant covers in a state is stronger than a cover which is identified with m relevant covers in that state if n is larger than m. In particular, a cover that has not been identified with any other is weak and normally not preferred as a domain of quantification.

What makes a concept stronger than another is not the semantic nature of the term which is used to denote it, e.g. proper names versus definite descriptions, but rather the number of alternative relevant concepts that have been identified with it in the specific information states of the relevant agents. In the example above, the concept 'John Smith' is stronger than the concept 'the upstairs neighbor' with respect to K's state, because K has identified the former concept with 'the major', whereas he has not identified 'the upstairs neighbor' with any other relevant concept. With respect to J's state the opposite holds. As a further illustration of the relativity of the notion of a strong concept, consider the concept of your mother. With respect to your information state such a concept is stronger than the concept 'the shortest spy' (unless you are different from most of us). On the other hand, in someone else's information state, say mine, the two concepts can instead be equally strong. This is due to the fact that in your information state the former concept is actually a *cluster* of concepts, but not in mine. Indeed, you can name your mother, you know what she looks like, and you can identify her by various means, whereas I cannot do anything of this sort. Thus, with respect to my information state, but not with respect to yours, the concept 'the reader's mother' will not be stronger than 'the shortest spy', which is also typically not identified with any other representation.

From this example it is clear that the objects of our experience are better represented by clusters of concepts, arising from the combination of elements of different covers, rather than by single isolated concepts. Strong concepts better represent genuine objects, than weak concepts and, therefore, the former can better serve as values for our variables. This is the meaning of STRENGTH.

Synopsis We have discussed four constraints so far. I suggest the following tentative ranking:

ANCHOR > *SHIFT, *ACCOMMODATION > STRENGTH

We have already seen that ANCHOR is assumed to be harder than *ACCOMMODATION in the literature (see Zeevat (1999a)). My hypothesis, which, however, needs further test, is that STRENGTH is weaker than *ACCOMMODATION. Indeed, we normally prefer weaker salient covers over stronger non salient ones. *SHIFT and *ACCOMMODATION are hard to compare. You cannot violate *SHIFT in order to satisfy *ACCOMMODATION and you cannot violate *ACCOMMODATION in order to satisfy *SHIFT. So I assume they are not ranked in any way.[13]

Putting aside the issue of the exact ordering between these constraints, notice that, by means of them, we can account for most cases of cover selection in ordinary situations. For instance, we can explain why demonstrative covers are normally preferred over alternative descriptive covers. First of all, if a demonstrative cover is available at all, then it is salient and, hence, its use cannot violate ANCHOR or *ACCOMMODATION. Furthermore, in standard situations, demonstrative covers are usually *clusters* of covers. If you have a person in front of you, you can point at her, but you can also describe her according to different parameters (visual image of her face, seize, age, etc.) and, in many cases, you can also name her. Demonstrative covers are normally strong and, hence, they

[13] This suggests that *SHIFT and *ACCOMMODATION could maybe be formulated together in a more general constraint. Notice that both are closely related to the principle Don't Overlook Anaphoric Possibilities (DOAP) from Williams (1997) also discussed in de Hoop and de Swart (1999) and de Hoop and Hendriks (1999), which requires to seize opportunities to anaphorize text. I leave this issue as a subject for future study.

are preferred also by STRENGTH. Note that typical cases in which demonstrative covers are not selected are cases like the man with the hood (see chapter 3), in which since the relevant people are dressed up, the demonstrative cover is crucially not identified with any good descriptive cover or with naming and, therefore, it is less strong.

The combination of the four constraints discussed above explains why in many ordinary cases we don't even notice the presence of conceptual covers and the classical notion of quantification seems to be sufficient. The addressee starts assuming the strongest contextually prominent cover or accommodates one, and just stays with it for the rest of the discourse. However we can easily find situations in which the weaker three constraints are crucially overruled and these are situations in which quantification under conceptual cover plays an essential role. I will discuss four cases that we have already encountered in the previous chapters, namely the double vision (chapter 2), the soccer game (chapter 3), the workshop (chapter 1), and the Ann-Bea (chapter 2) situations. In these examples, *SHIFT is overruled in order to avoid violations of general principles of rational conversation.[14] The suggested ranking is then that the latter are harder than the former. This is in accordance with what is normally assumed in the literature (see Zeevat (1999a), and also Stalnaker (1978)). I will discuss three of such general principles of conversation: CONSISTENCY, *TRIVIAL and RELEVANCE.

CONSISTENCY is a constraint which expresses preference for interpretations that do not conflict with the context (see Grice's Maxim of Quality, Stalnaker's first principle of rational conversation in (1979), and also Zeevat (1999a) and van der Sandt (1992)). Recall the following situations discussed in chapter 2 and 3 respectively, in which CONSISTENCY plays a crucial role:

double vision Ralph ascribes contradictory properties to Ortcutt since, having met him on two quite different occasions, he is 'acquainted' with him in two different ways.

(214) a. Ralph believes Ortcutt to be a spy; and Ralph believes Ortcutt not to be a spy.

b. $\exists x_n(x_n = o \wedge \Box S(x_n)) \wedge \exists y_m(y_m = o \wedge \Box\neg S(y_m))$

the soccer game Suppose you are attending a soccer game. All of the 22 players are in your perceptual field. You know their names, say a, b, c, ..., but you don't recognize any of them. Consider the following sentence:

(215) a. Anyone might be anyone.

[14] See Grice's theory of conversation and Stalnaker (1978).

b. $\forall x_n \forall y_m \Diamond (x_n = y_m)$

In both the double vision and the soccer game examples we have two conflicting constraints. On the one hand, we have *SHIFT which suggests to interpret m as n. On the other, the fulfillment of CONSISTENCY prevents this resolution, because it would render the sentences inconsistent with the context. In the double vision case, this resolution leads to inconsistency, because the sentence would say that Ralph's beliefs are contradictory, but we know that this is not the case as Ralph is not logically insane, since he simply lacks some information. In the soccer game, if n and m are assigned one and the same value the sentence is false in the described situation in which the relevant domain is not a singleton. Intuitively, an assignment of different values for n and m is what is normally assumed by an interpreter of such sentences, this suggests that CONSISTENCY is harder than *SHIFT and confirms what is normally assumed in the literature.

***TRIVIAL** is a constraint which forbids under-informative interpretations (see Stalnaker's first principle of rational conversation). Recall the original workshop example discussed in chapter 1 in which *TRIVIAL conflicts with *SHIFT.

the workshop You are attending a workshop. In front of you lies the list of names of all participants, around you are sitting the participants in flesh and blood. Consider the following dialogue, question, and assertion, uttered in such a situation:

(216) A: Who$_n$ is that man?

B: That man is Ken Parker.

A: Who$_m$ is Nathan Never?

B: Nathan Never is the one over there.

(217) Who$_n$ is who$_m$?

(218) I don't know who$_n$ is who$_m$.

Again if we follow *SHIFT, then n and m should be assigned the same value. On the other hand, if *SHIFT is satisfied, *TRIVIAL which forbids trivial interpretations, would be violated. Since, intuitively, we shift conceptualization while interpreting these sentences, the example suggests that *TRIVIAL ranks higher than *SHIFT.

RELEVANCE expresses preference for relevant interpretations (see Grice's maxim of relation and Horn's I-principle). A formal characterization of the notion of relevance has been recently attempted by a number of authors, for instance Roberts (1996b), Groenendijk (1999), and in particular van Rooy (2000a), who also discusses examples showing how relevance can influence the interpretation of belief attributions.[15] I will not discuss these formalizations though, because the intuitive notion of relevance which Grice had in mind, suffices to make my point here. Consider the following slight modification of the Ann-Bea example we discussed in chapter 2.

Ann and Bea In front of Ralph stand two women. For some reason we don't need to investigate, Ralph believes that the woman on the left, who is smiling, is Bea and the woman on the right, who is frowning, is Ann. As a matter of fact, exactly the opposite is the case. Bea is frowning on the right and Ann is smiling on the left. In the picture w_1 is the world of evaluation and w_2 is the only possibility in Ralph's belief state.

$$\begin{array}{lccc} w_1 & \mapsto & (☺) & (☹) \\ & & [ann] & [bea] \\ \\ w_2 & \mapsto & (☺) & (☹) \\ & & [bea] & [ann] \end{array}$$

There are two possible conceptual covers in such a situation, namely the set A=$\{\lambda w[\text{left}]_w,\lambda w[\text{right}]_w\}$, which cross-identifies the women which stand in the same perceptual relation to Ralph and the set B=$\{\lambda w[\text{Ann}]_w, \lambda w[\text{Bea}]_w\}$, which cross-identifies the women by their name. Suppose all of a sudden Ralph starts chasing the woman on the left to bring her to a mental institution. I ask you: 'Why is Ralph chasing Ann?'. You answer:

(219) Ralph believes that Ann is insane.

There are three possible ways of interpreting this sentences in the described situation: (a) an interpretation *de re* under cover A, in which Ann is identified as the woman on the left; (b) an interpretation *de re* under cover B, in which Ann is identified as Ann; (c) the *de dicto* interpretation.

All three interpretations are consistent with the background. Interpretation (a) seems to involve a violation of *SHIFT, and probably STRENGTH. Indeed, my question, which explicitly uses 'Ann' to identify the relevant woman, suggests cover B as the prominent one. (b) and (c) do not involve such violation(s). Still, intuitively, we prefer interpretation (a) for (219) in such situation. I suggest

[15]The Ann-Bea case I discuss here has been inspired by van Rooy's example of the English gentleman.

that the reason is that only under such an interpretation the sentence would be relevant. Indeed, whether the belief attribution (219) is contributing to explain for us Ralph's behaviour depends on how Ann is identified in Ralph's belief state. Whether or not Ralph believed that Ann – who is, according to him, the woman on the right– is insane does not help explaining why he is chasing the woman on the left, whereas the fact that he believes that the woman on the left is insane does contribute to an explanation. Thus, only under interpretation (a) the belief attribution constitutes a proper answer to my question and hence is relevant. This is why the addressee adopts a cover containing the concept 'the woman on the left' in such a situation, although this involves a violation of *SHIFT. By assuming that RELEVANCE is harder than *SHIFT we can account for this intuition.

Synopsis

We have discussed the following interpretation constraints which seem to play a role in the operation of cover selection:

ANCHOR Interpretation should be anchored to the context.

CONVERSATIONAL MAXIMS

> **CONSISTENCY** avoid inconsistent interpretations;
>
> ***TRIVIAL** avoid trivial interpretations;
>
> **RELEVANCE** prefer relevant interpretations.

***SHIFT** Do not shift conceptualization.

***ACCOMMODATION** Do not accommodate.

STRENGTH Stronger covers are preferred.

The following is a possible ranking consistent with the phenomena discussed above:

ANCHOR, C. MAXIMS > *SHIFT, *ACCOMMODATION > STRENGTH

If we order the principles we discussed in this way,[16] we are able to explain the process of saturating CC-indices in many of the cases we have discussed in the thesis. As a further illustration I will consider the case of Susan's mother, discussed in chapter 2. The interpretation procedure in this example involves a number of contrasting constraints and clearly illustrates the nature of an optimality theoretic explanation.

[16]By the assumption of other ordered constraints we might have reached exactly the same result. This is not really important here, what is relevant is that constraints play a role here which can be overruled by higher ranking ones.

Susan's mother For ease of reference, I rewrite the relevant situation described by van Fraassen (1979):

> Susan's mother is a successful artist. Susan goes to college, where she discusses with the registrar the impact of the raise in tuition on her personal finances. She reports to her mother 'He said that I should ask for a larger allowance from home'. Susan's mother exclaims:
>
> (220) He must think I am rich.
>
> Susan, looking puzzled, says 'I don't think he has any idea who you are'.[17]

van Fraassen analyzes the example as follows:

> The information the mother intends to convey is that the registrar believes that Susan's mother is rich, while Susan misunderstands her as saying that the registrar thinks that such and such successful artist is rich. The misunderstanding disappears if the mother gives information about herself, that is, about what she had in mind. She relied, it seems, on the auxiliary assertion 'I am your mother'.[18]

A number of conflicting constraints play a role in the interpretation of (220). On the one hand, we have *SHIFT, *ACCOMMODATION and STRENGTH which forbid the selection of a cover containing the concept 'Susan's mother', because it is not prominent and clearly weaker than other salient ones with respect to the registrar's information state. On the other hand, we have CONSISTENCY which, if fulfilled, forces precisely such a non-salient and weak interpretation. Susan accepts the sentence after the clarification of the mother, and this confirms our hypothesis that CONSISTENCY can overrule *SHIFT, *ACCOMMODATION and STRENGTH.

Suppose now that Susan's mother utters the following sentence in the same situation:

(221) He must think I am your mother.

Only under a conceptualization containing the concept 'Susan's mother', the belief attribution in (221) is consistent with the common ground, but, under such a cover, the sentence is clearly trivial. Thus, the only way to satisfy CONSISTENCY here would involve a violation of *TRIVIAL. This explains why the sentence is pragmatically unacceptable in such a situation. Still, it seems that the inconsistent reading is preferred over the trivial one in this case and the present analysis can explain this fact as follows. While the inconsistent reading

[17] van Fraassen (1979), pp. 371-372 and chapter 2.
[18] van Fraassen (1979), p. 372.

just violates CONSISTENCY, the trivial reading violates *TRIVIAL, but also *SHIFT, *ACCOMMODATION and STRENGTH, and these violations of lower constraints become crucial in this case.

The following two diagrams summarize our OT-analysis of the case of Susan's mother. I use (*) to indicate that the interpretation violates the corresponding constraint, and !(*) to indicate a crucial violation. Optimal interpretations are those which do not involve any crucial violation.

(220) He must think I am rich.

(a) *de re* interpretation under the prominent cover containing the concept 'such and such successful artist': $\lambda x \Box R(x)(a)$;

(b) *de re* interpretation under a cover containing the concept 'Susan's mother': $\lambda x \Box R(x)(m)$.

(220)	***TRIV**,	**CONS**	***SHIFT**,	***ACC**	**STRENGTH**
(a)		!(*)			
(b)			(*)	(*)	(*)

(221) He must think I am your mother.

(a) *de re* interpretation under the prominent cover containing the concept 'such and such successful artist': $\lambda x \Box M(x)(a)$

(b) *de re* interpretation under a cover containing the concept 'Susan's mother': $\lambda x \Box M(x)(m)$

(221)	***TRIV**,	**CONS**	***SHIFT**,	***ACC**	**STRENGTH**
(a)		(*)			
(b)	(*)		!(*)	!(*)	(*)

We will return to this example later on.

Bi-Dimensional Optimality Theory

The OT analysis discussed so far enables an explanation of the process of saturating CC-indices in many of the cases we have discussed in the thesis, but not all. As an illustration consider the following example.

the bald president Consider the following situation. Naming is the prominent cover and the addressee holds as common ground that: (i) Putin is the actual president of Russia; (ii) Ralph believes that Jeltsin is the actual president of Russia; (iii) Ralph would not assent to the sentence: 'Putin is bald'. Consider now the following sentence uttered in such a situation:

(222) Ralph believes that Putin is bald.

Let A be naming and B be a cover containing the concept 'the actual president of Russia'. The sentence has three possible interpretations in such circumstances: (a) the *de dicto* reading; (b) the *de re* reading under A; and (c) the *de re* reading under B. According to the interpretation theory I have discussed so far, interpretation (c) would be optimal given the situation. Indeed, although such an interpretation violates *SHIFT, *ACCOMMODATION (and probably STRENGTH), the other two alternative interpretations, which satisfy these constraints, crucially violate CONSISTENCY, because of clause (iii) above.

(222)	**CONS**	***SHIFT**,	***ACC**	**STRENGTH**
(a) *de dicto*	!(*)			
(b) *de re*-naming	!(*)			
(c) *de re*-descript.		(*)	(*)	(*)

As illustrated by the constraint table, interpretation (c) is the predicted optimal interpretation. On such interpretation, the sentence says that (Putin is the actual president of Russia and) Ralph would assent to the sentence 'The actual president of Russia is bald'. Since Ralph believes that Jeltsin is the actual president of Russia, (c) also entails the *de dicto* reading of the sentence: 'Ralph believes that Jeltsin is bald'. This prediction is clearly counter-intuitive. An intuitive explanation of why reading (c) is not preferred in such a situation is that a speaker expressing such a content by means of such a sentence would not be cooperative. Indeed, in the described situation, the same content could have been conveyed in a much more efficient way by uttering the following sentence:

(223) Ralph believes that the president of Russia is bald.

The *de dicto* reading of this alternative formulation and reading (c) of (222) convey the same information in the described situation in which the information that Putin is the actual president of Russia is part of the common ground. But the former interpretation does not involve any shift of cover or accommodation. For this reason, (223) is more efficient than (222), and, therefore, the speaker, if cooperative, should have chosen it. This is Grice's principle of cooperation. A speaker has a responsibility of what the audience will make of her sentences. In cooperative exchanges, she goes through the interpretation herself and makes sure that the intended content is as easy to obtain as possible. A cooperative speaker would never have uttered (222) to convey the information that Ralph would assent to the sentence 'Jeltsin is bald'. Therefore an interpretation of (222) which conveys such information cannot be optimal in such a situation. Note, however, that such an explanation cannot be formulated in the OT interpretation theory we have considered so far, in which inputs are given by single sentences and no reference is made to alternative sentences that the speaker might have used. In order to account for these cases, we need a more complex analysis, where the optimal solution is searched on two dimensions, rather than one: the dimension of the addressee and the one of the speaker, and in which the two optimization procedures of the addressee and of the speaker can refer to each other and crucially constrain each other. Such an analysis is the bi-directional Optimality Theory of Reinhard Blutner (see Blutner (1999) and Blutner and Jäger (1999)). In the next section, I follow Dekker and van Rooy (1999) and define bi-directional OT interpretation as an 'interpretation game'. The use of game-theoretical concepts allows a perspicuous formulation of Blutner's central notions of strong and weak optimality. Furthermore, a game-theoretic formulation of the process of interpretation will be useful in order to account for the interplay between the addressee and the speaker with their particular interests and goals in the interpretation of context dependent natural language expressions.

4.3.2 Interpretations as Games

In an interesting recent article,[19] Dekker and van Rooy (D&vR) propose to apply concepts that have been studied in the field of Game Theory[20] to investigate a series of phenomena in the semantic/pragmatic interface. They rewrite OT interpretation theory in terms of game-theoretic notions where optimality itself is viewed as a solution concept for a game.

The central notion introduced by D&vR is that of an interpretation game. Interpretation games are defined in terms of 'strategic games'.

A *strategic game* G is a triple

$$G = (N, (A_i)_{i \in N}, (\succ_i)_{i \in n})$$

where N is a set of players, $(A_i)_{i \in N}$ maps each player to a non-empty set of alternative actions A_i, and $\succ_i$ is a preference relation for player i over the product $A = A_1 \times \ldots \times A_n$ of possible actions of all players. An element a of such a product A is called an action *profile*.

An *interpretation game* I is a strategic game involving two players, the Speaker and the Hearer, $N = \{\text{S, H}\}$. The set of alternative actions for the speaker consists of a set $A_S = \{F1, F2, ...\}$ of possible *forms*, the set of alternative actions for the hearer consists of a set $A_H = \{C1, C2, ...\}$ of possible *contents*. S chooses a suitable form $F \in A_S$ for a content $C \in A_H$ to be communicated. H chooses a suitable interpretation $C \in A_H$ for a signaled representation $F \in A_S$. Optimality theoretic preferences are used in combination with particular goal-directed preferences to define the preference relations of the speaker $\succ_S$ and of the hearer $\succ_H$. The relations $\succ_S$ and $\succ_H$ should be interpreted as strict preferences and hence are taken to be transitive, anti-reflexive and anti-symmetric.

The interpretation games I will consider are crucially played in a specific context. I will therefore identify forms with utterances and interpretations with

[19]Dekker and van Rooy (1999).

[20]See Osborne and Rubistein (1994) for an introduction to the main game-theoretical concepts.

actions on the specific information state which constitutes the common-ground in the circumstances of these utterances (see Blutner (1999) who makes the same assumption). The fulfillment of the interpretation constraints discussed in the previous section and general principles of cooperation will determine the preference relation of the hearer $\succ_H$. General principles of generation, cooperativity and the particular goals of the speaker will interplay in the determination of $\succ_S$. These preference relations are highly context dependent. In the present analysis, they are sensitive to three specific aspects of the context: (a) which covers are salient; (b) which information is presupposed by the speaker and by the addressee; (c) the specific intentions of the speaker. The first two factors are relevant in that they determine whether or not a profile satisfies the interpretation constraints discussed above, and hence influence the preference relation of the addressee and of the speaker, if cooperative; the third factor helps in determining which content is intended by the speaker who has authority on how her utterance should be interpreted, and hence influences the preference relation of the speaker.[21]

One of the tasks of game theory is to develop solution concepts which allow one to make predictions about the outcome of a game and about how the players will interact. In D&vR, optimality is viewed as a solution concept of an interpretation game. Optimal solutions are no longer optimal interpretations of a given expression, but optimal profiles consisting of an utterance and an interpretation. D&vR discuss two notions of optimality: the notion of *Nash-optimality*, and *BJ-optimality*. The first notion is nothing else than the well-known solution concept of Nash equilibrium, which is shown to be the game-theoretic equivalent of Blutner's (1999) notion of strong optimality. BJ-optimality is the game-theoretic counterpart of Blutner's notion of weak optimality. I will present D&vR's definitions of these two notions. I will then show how these two solution concepts can be used to account for our intuitions about the bald president example discussed above.

Nash- and BJ-optimality

In Blutner's (1999) bi-directional OT, a mechanism compares different possible interpretations C for the same syntactic expression F and another mechanism compares different possible syntactic formulations F for the same content C. A form-content pair (F,C) is then *strongly optimal* just in case C is an optimal interpretation for F according to the first mechanism and F is an optimal form

[21] Eventually the specific intentions of the speaker might also influence the preference relation of the addressee, if cooperative. This might be the sense of cooperativity from the addressee point of view. Being cooperative means to minimize effort for the other participant. S is cooperative, if she chooses the form which can be interpreted in the most straightforward way by H. H cannot do much to help S in choosing the right words, but she can be cooperative by selecting the intended interpretation. Obviously only correctly informed agents can be cooperative. If S and H fail to share a common-ground or if H fails to know the intentions of S, misunderstanding can arise and communication can break down.

for C according to the second mechanism. D&vR have shown that this notion of strong optimality can be perspicuously formalized by means of the classical solution concept of a Nash equilibrium. Blutner's strong optimal solutions are identified with Nash equilibria in an interpretation game.

Given an action profile $a \in A$ and an action $a_i \in A_i$, let $a[i : a_i]$ denote the profile which is like a, but with player i taking action a_i.

4.3.4. Definition. [Nash-optimality] Let $I = (N, (A_i)_{i \in N}, (\succ_i)_{i \in n})$ be an interpretation game. An action profile a is *Nash-optimal* in I, $\text{NASH}_I(\text{a})$ iff

$$\forall i \in N: \; \forall a_i \in A_i : \neg(a[i : a_i] \succ_i a)$$

Intuitively, in a Nash equilibrium, every player acts optimally given the other players' actions, that is, every player's action is the best response to the choices of the other players.

As an illustration consider the interpretation game depicted by means of the following matrix:

	C1	*C2*
F1	$(2, 3)$	$(4, 5)$
F2	$(3, 2)$	$(1, 1)$

In such matrices, the Speaker chooses the row and the Hearer the column to be played and preference relations are formulated in terms of payoff functions,[22] where the payoff pair (x, y) expresses that the speaker gets payoff x and the hearer gets payoff y. In the game depicted by this specific matrix, S prefers causing $C1$ by performing $F2$, whereas $F1$ is the preferred way of causing $C2$. H prefers performing action $C2$ as a response to $F1$, and $C1$ as a response to $F2$. The game has two Nash equilibria, namely the profiles $(F2, C1)$ and $(F1, C2)$.

In the definition of a Nash equilibrium the only strict preferences $\succ_i$ which really count are those between two profiles a and b if their only difference lies in the choice of $i \in \{S, H\}$, i.e. if $a = b[i : b_i]$ for some b_i. For this reason D&vR propose to represent Nash equilibria in interpretation games[23] by drawing arrows between two profiles on the same row or in the same column, with the following meaning: $\leftarrow$ means 'H strictly prefers the left profile', $\rightarrow$ means 'H strictly prefers the right profile', $\downarrow$ 'S strictly prefers the bottom profile', and $\uparrow$ 'S strictly prefers the top profile'. The game above is then represented by the following table in which the Nash equilibria are immediately visualized by $\circ$:

$$\begin{array}{c|ccc|}
 & \mathit{C1} & & \mathit{C2} \\
\hline
\mathit{F1} & & \rightarrow & \circ \\
 & \downarrow & & \uparrow \\
\mathit{F2} & \circ & \leftarrow & \\
\hline
\end{array}$$

[22] Preference relations can be expressed in terms of payoff functions $(u_i)_{i \in N}$, where $u_i : A \rightarrow R$ is the payoff function of player i. Action profiles with higher payoff are preferred.

[23] Interpretation games crucially involve only two players.

If no arrow is leaving from a profile a, then a is a Nash equilibrium. This means that a profile (F, C) is Nash-optimal in I iff for all contents $CN \in A_H$ and forms $FN \in A_S$ in I:

(i) $(F, CN) \not\succ_H (F, C)$

(ii) $(FN, C) \not\succ_S (F, C)$

By means of the notion of Nash-optimality, we can characterize anomalous interpretations. Intuitively, a pair (F, C) is anomalous with respect to I iff it is *not* Nash-optimal in I, and this is the case iff either C is not an optimal interpretation for F in I (clause (i) is not satisfied) or, if C is an optimal interpretation, then C could have been expressed more efficiently by an alternative form (clause (ii) is not satisfied).

The strong version of optimality characterized by the notion of a Nash equilibrium is useful to explain many standard cases, but it has been shown not to be always satisfactory. In his (1999) article, Blutner illustrates this by means of the following example inspired by Horn:

(224) Black Bart killed the sheriff.

(225) Black Bart caused the sheriff to die.

The lexical causative *kill* tends to be restricted to stereotypical causative situations (e.g. Black Bart shot the sheriff), and the marked construction in (225) tends to refer to more marked situations (e.g. Black Bart caused the sheriff's gun to backfire by stuffing it with cotton). The general tendency illustrated by this example seems to be that 'unmarked forms tend to be used for unmarked situations and marked forms for marked situations' (Horn (1984), p. 26). This tendency has been called by Horn *the division of pragmatic labour*.

This case can be formalized by means of the following interpretation game:

$$
\begin{array}{r|ccc|}
 & \mathit{C1} & & \mathit{C2} \\
\hline
\mathit{F1} & & \rightarrow & \\
 & \downarrow & & \downarrow \\
\mathit{F2} & & \rightarrow & \circ \\
\hline
\end{array}
$$

where $F1$ and $F2$ stand for the marked and the unmarked forms respectively and $C1$ and $C2$ stand for the the marked and the unmarked situation respectively. By the notion of Nash-optimality we can account for the fact that (224) picks up stereotypical situations. Unmarked forms ($F2$) are preferred over marked forms ($F1$), and stereotypical situations ($C2$) are easier to understand than atypical situations ($C1$). The profile unmarked form-unmarked situation $(F2, C2)$ is Nash in such a game. But Nash-optimality is not sufficient to explain why (225) obtains the unusual interpretation. Indeed, no interpretation is selected for the marked form $F1$. The profile marked form-marked content is intuitively chosen because (i) the alternative unmarked form does not get the marked interpretation and (ii) we prefer to use the unmarked form to express the unmarked situation. Now, by means of the notion of optimality defined in terms of a Nash equilibrium we cannot capture this kind of reasoning. A profile is Nash-optimal iff it is optimal for the Speaker and optimal for Hearer and these two checks for optimality are independent of each other. The search for the optimal choice for one player is not influenced by the preference relation of the other player. In order to account for H's reasoning in this case, we need a notion in which the two optimization procedures of the hearer and of the speaker can refer to each other and constrain each other. Such a notion is the notion of *weak optimality* introduced in Blutner (1998). D&vR's notion of BJ-optimality is the perspicuous game-theoretical formulation of such notion.

BJ-optimality is defined as follows (see Dekker and van Rooy (1999) for further discussion):

4.3.5. Definition. [BJ-Optimality] Let $I = (N, (A_i)_{i\in N}, (\succ_i)_{i\in N})$ be an interpretation game. Then the set BJ_I of BJ-optimal solutions in I is defined as follows:

$$BJ_I = NASH_{I_n}$$

where I_n is the fixed point, i.e. $I_{n+1} = I_n$, of the sequence of games $I_0, \ldots I_m, \ldots$ constructed as follows:

(i) $I_0 = I$

(ii) $I_{n+1} = (N, (A)_{i\in N}, (\succ_{i_{n+1}})_{i\in N})$ with

(a) $\succ_{S_{n+1}} \; = \; \succ_{S_n} \setminus \{(y, z) \mid \exists x \in NASH_{I_n} : x \succ_{H_n} y\}$;

(b) $\succ_{H_{n+1}} \; = \; \succ_{H_n} \setminus \{(y, z) \mid \exists x \in NASH_{I_n} : x \succ_{S_n} y\}$.

In the construction of I_{n+1} you eliminate preferences for blocked profiles y. A profile y is blocked in I_{n+1}, if there was a Nash-optimal profile x which was preferred to y in I_n. If $I_{n+1} = I_n$, then the Nash equilibria of I_n are the BJ-optimal solutions in I_0. That is, if an action profile a is a Nash-optimal solution in the fixed point game of the sequence generated from a game I, then a is BJ-optimal in I.

The intuitive idea of this construction is that Nash-optimal profiles block less preferred ones and preferences for blocked profiles are overruled. As an illustration, let's go back to the game I determined by Horn's sheriff example. The sequence generated from such game consists of the two games represented in the following matrices where blocked profiles are indicated by $\perp$:

$$I_0: \quad \begin{array}{c|ccc|} & \mathit{C1} & & \mathit{C2} \\ \hline \mathit{F1} & & \rightarrow & \bot \\ & \downarrow & & \downarrow \\ \mathit{F2} & \bot & \rightarrow & \circ \\ \hline \end{array} \qquad\qquad I_1: \quad \begin{array}{c|ccc|} & \mathit{C1} & & \mathit{C2} \\ \hline \mathit{F1} & \circ & & \bot \\ & & & \downarrow \\ \mathit{F2} & \bot & \rightarrow & \circ \\ \hline \end{array}$$

I_0 is I, and I_1 is obtained from I_0 by eliminating preferences for the two blocked profiles $(F2, C1)$ and $(F1, C2)$. I_1 has two Nash-optimal solutions: $(F1, C1)$ and $(F2, C2)$. Since no preference can be eliminated in the next step of our construction (i.e. $I_1 = I_2$), these two profiles are the two BJ-optimal solutions of I.

$$\begin{array}{c|ccc|} & \mathit{C1} & & \mathit{C2} \\ \hline \mathit{F1} & \text{BJ} & \rightarrow & \bot \\ & \downarrow & & \downarrow \\ \mathit{F2} & \bot & \rightarrow & \circ \\ \hline \end{array}$$

The profiles $(F1, C2)$ and $(F2, C1)$ are not BJ-optimal in I, because overruled by the Nash-optimal $(F2, C2)$. On the other hand, $(F1, C1)$ is BJ-optimal. Although $(F1, C2) \succ_H (F1, C1)$ and $(F2, C1) \succ_S (F1, C1)$, these preference do not count because $(F1, C2)$ and $(F2, C1)$ are blocked by the Nash-optimal $(F2, C2)$, since $(F2, C2) \succ_S (F1, C2)$ and $(F2, C2) \succ_H (F2, C1)$. In the notion of BJ-optimality, a player's perspective on optimization is crucially constrained by the other player's perspective and *vice versa*. We thus capture H's intuitive reasoning in the sheriff case. The profile marked form-marked content $(F1, C1)$ is intuitively chosen because (i) the alternative unmarked form $F2$ does not get the marked interpretation $C1$ ($(F2, C2) \succ_H (F2, C1)$) and (ii) we prefer to use the unmarked form $F2$ to express the unmarked situation $C2$ ($(F2, C2) \succ_S (F1, C2)$). The hearer chooses the marked $C1$ rather than the unmarked $C2$ as interpretation for $F1$, because she can reason as follows: if the speaker had wanted to communicate $C2$ he would have chosen the Nash-optimal $F2$.

Now let us see how the bald president case discussed above can be accounted for by means of these notions.

the bald president For ease of reference, I restate the situation. Naming is the prominent cover. The common ground contains the following information: (i) Putin is the actual president of Russia; (ii) Ralph believes that Jeltsin is the actual president of Russia; (iii) Ralph would not assent to the sentence: 'Putin is bald'. In such context, the following sentence is uttered by the speaker S:

(222) Ralph believes that Putin is bald.

Intuitively, a rational addressee H can do two things in such a situation: either refute to perform any action or consider revising her state with the information that Ralph does believe *de dicto* that Putin is bald. In any case, H does not update with the information that Ralph would assent to the sentence: 'Jeltsin is bald'. This last action was predicted as optimal by the one-dimensional OT interpretation theory I introduced in the previous section. Let's see whether the two-dimensional theory I have just described does any better here.

I propose to characterize such a situation by means of the following interpretation game:

$$\begin{array}{c|ccc|}
 & \mathit{C1} & & \mathit{C2} \\ \cline{2-4}
\mathit{F1} & & \rightarrow & \\
 & \uparrow & & \downarrow \\
\mathit{F2} & & \rightarrow & \\ \cline{2-4}
\end{array}$$

$F1$ and $F2$ are the utterances of the sentences:

(222) Ralph believes that Putin is bald.

(223) Ralph believes that the actual president of Russia is bald.

For ease of reference, I will denote the first action by '*Putin*' and the second action by '*the president*'.

$F1 =$ '*Putin*'

$F2 =$ '*the president*'

Let us see now how $C1$ and $C2$ are characterized. Let A be naming and B be a cover containing the concept 'the actual president of Russia'. Assume A and B are the only two covers available in our situation. Each of the two sentences above has then three possible interpretations. Let $\wp$ be a conceptual perspective such that $\wp(n) = A$ and $\wp(m) = B$.

(222) Ralph believes that Putin is bald.

a. *de dicto*: $\Box B(p)$

b. *de re* under A: $\exists x_n[x_n = p \wedge \Box B(x_n)]$

c. *de re* under B: $\exists x_m[x_m = p \wedge \Box B(x_m)]$

(223) Ralph believes that the actual president of Russia is bald.

d. *de dicto*: $\Box B(r)$

e. *de re* under A: $\exists x_n[x_n = r \wedge \Box B(x_n)]$

f. *de re* under B: $\exists x_m[x_m = r \wedge \Box B(x_m)]$

Given our characterization of the situation, these six possible interpretations collapse in only two different possible actions on the relevant common ground. Let σ stand for the common ground in the described situation and let $\wp$ be as above, we then obtain the following equivalences:[24]

[24]Recall that σ supports the information that Putin is the actual president of Russia, $r = p$.

(α) $UP^{\wp}_{\sigma}(a) = UP^{\wp}_{\sigma}(b) = UP^{\wp}_{\sigma}(e)$

(β) $UP^{\wp}_{\sigma}(c) = UP^{\wp}_{\sigma}(d) = UP^{\wp}_{\sigma}(f)$

The updates of σ with the six interpretations above collapse in only two possible updates. There are only two possible alternative actions for H on the relevant state: the action in (α) which consists in eliminating those possibilities in σ in which it is true that Ralph believes *de dicto* that Putin is bald; the action in (β) which consists in eliminating those possibilities in σ in which it is true that Ralph believes *de dicto* that the actual predident of Russia is bald. For ease of reference, I will denote the first action by $UP(put)$ and the second action by $UP(pres)$. I will identify $C1$ and $C2$ with these two possible updates:

$$C1 = UP(put)$$

$$C2 = UP(pres)$$

Let us turn now to the preference relations $\succ_H$ and $\succ_S$. H's preferences are obtained by the following two OT-constraint tables for the two relevant utterances and contents:

F1	**CONS**	***SHIFT**,	***ACC**	**STRENGTH**
C1	!(*)			
C2		(*)	(*)	(*)

$F2$	**CONS**	***SHIFT**,	***ACC**	**STRENGTH**
$C1$	!(*)			
$C2$				

$C1$ crucially violates CONSISTENCY in both diagrams because such an action on the common ground leads to the absurd state in any case. Therefore, in our game, H strictly prefers a profile $(FN, C2)$ over $(FN, C1)$. Consistent interpretations are preferred over inconsistent interpretations.

As for the speaker's preferences, I assume in the matrix that S strictly prefers causing $C1$ by performing $F1$ and causing $C2$ by performing $F2$. Indeed, an utterance of the sentence: 'Ralph believes that Putin is bald' (i.e. $F1$) is the most cooperative way of conveying the information that Ralph would assent to the sentence: 'Putin is bald' (i.e. the information brought about by $C1$) and an utterance of the sentence: 'Ralph believes that the actual president of Russia is bald' (i.e. $F2$) is the most cooperative way of conveying the information that Ralph would assent to the sentence: 'The actual president of Russia is bald' (i.e. the information brought about by $C2$). Therefore, if we assume that S is cooperative, we obtain the following preference relation $\succ_S$:

$$(\text{'}\mathit{Putin}\text{'}, UP(put)) \succ_S (\text{'}\mathit{the\ president}\text{'}, UP(put))$$

$$(\text{‘}\textit{the president}\text{’}, UP(pres)) \succ_S (\text{‘}\textit{Putin}\text{’}, UP(pres))$$

Cooperative formulations are preferred over non-cooperative formulations. While the first preference does not have any effect on the following discussion, the second preference is crucial to the outcome of the game.[25]

Our game has one Nash equilibrium, namely the profile (‘*the president*’, $UP(pres)$):

$$\begin{array}{r|ccc|}
 & UP(put) & & UP(pres) \\
\hline
\text{‘}\textit{Putin}\text{’} & & \rightarrow & \\
 & \uparrow & & \downarrow \\
\text{‘}\textit{the president}\text{’} & & \rightarrow & \circ \\
\hline
\end{array}$$

Nash-optimality selects $UP(pres)$ as optimal interpretation for ‘*the president*’ because consistent interpretations are preferred over inconsistent interpretations and efficient formulations are preferred over non-efficient formulations. But Nash-optimality does not select any interpretation for ‘*Putin*’. Its interpretation is left open because (i) profile (‘*Putin*’,$UP(put)$) is anomalous since $UP(put)$ is not an optimal interpretation (it leads to inconsistency) and (ii) (‘*Putin*’,$UP(pres)$) is anomalous since although $UP(pres)$ is an optimal interpretation, it could have been expressed more efficiently by an alternative form. In order to account for the fact that the inconsistent interpretation of (222) under the prominent cover is preferred by the hearer over an interpretation under the problematic conceptualization we need the weaker notion of BJ-optimality. Intuitively the hearer chooses action $UP(put)$, which would lead her to the absurd state, rather than $UP(pres)$ as a response to ‘*Putin*’ because she can reason as follows: If the Speaker had wanted to convey the consistent interpretation $UP(pres)$, then S should have chosen the more efficient formulation ‘*the president*’. But S chose ‘*Putin*’. Thus S must have meant to convey $UP(put)$. This is precisely the kind of reasoning captured by the notion of BJ-optimality. Indeed, profile (‘*Putin*’,$UP(put)$) is BJ-optimal in our game. Whereas (‘*Putin*’,$UP(pres)$) is not, because overruled by the Nash-optimal (‘*the president*’,$UP(pres)$).

$$\begin{array}{r|ccc|}
 & UP(put) & & UP(pres) \\
\hline
\text{‘}\textit{Putin}\text{’} & \text{BJ} & \rightarrow & \perp \\
 & \uparrow & & \downarrow \\
\text{‘}\textit{the president}\text{’} & \perp & \rightarrow & \circ \\
\hline
\end{array}$$

[25] This preference is further justified by the constraint table above. Indeed, $C2$ violates *SHIFT, *ACC and STRENGTH in the first diagram, but not in the second. This is due to the fact that according to our semantics $F1$ can produce in σ the update effect $C2$ only by means of an update with reading (c), which, as we have seen, involves such violations. Whereas $F2$ can convey such information by means of reading (d), which does not involve any constraint violation. Therefore, the speaker, who is assumed to be cooperative, crucially prefers profile $(F2, C2)$ over $(F1, C2)$.

We can now explain the addressee's behaviour in our presidential example. In order to interpret an utterance of 'Ralph believes that Putin is bald', the addressee does not adopt a cover containing the concept 'the actual president of Russia' (profile ('*Putin*',$UP(pres)$) is blocked), but she rather assumes the prominent conceptualization (profile ('*Putin*',$UP(put)$) is BJ-optimal). The latter action leads her to the absurd state. She can protest or she can decide to start a process of revision of her information.

How is the bald president situation different from the case of Susan's mother that we have considered above? As in the presidential example, in the case of Susan's mother we have a sentence whose only interpretation which does not contradict the common ground involves a violation of *SHIFT, *ACCOMMODATION and STRENGTH. In the presidential example, such an interpretation was intuitively unacceptable and we could account for it. Instead, in the case of Susan's mother, such an interpretation is intuitively acceptable. We have to show how it is possible that the violation of exactly the same interpretation principles can be tolerated on one occasion and not on the other. Although the two cases have a similar structure, there is a crucial difference. In the latter case, we can find a series of reasons justifying the Speaker's chosen utterance, and hence, although the chosen formulation of the intended content is clearly uncooperative, it is hard to find an alternative formulation which is *strictly* preferred by S over the one actually used. This, I suggest, explains our different intuitions about the addressee reactions in the two cases. Let us have a closer look.

Susan's mother again For ease of reference, I restate the situation:

> Susan's mother is a successful artist. Susan goes to college, where she discusses with the registrar the impact of the raise in tuition on her personal finances. She reports to her mother 'He said that I should ask for a larger allowance from home'. Susan's mother exclaims: 'He must think I am rich.' Susan, looking puzzled, says 'I don't think he has any idea who you are'.[26]

van Fraassen analyzes the example as follows:

> The information the mother intends to convey is that the registrar believes that Susan's mother is rich, while Susan misunderstands her as saying that the registrar thinks that such and such successful artist is rich. The misunderstanding disappears if the mother gives information about herself, that is, about what she had in mind. She relied, it seems, on the auxiliary assertion 'I am your mother'.[27]

[26] van Fraassen (1979), pp. 371-372 and chapter 2.
[27] van Fraassen (1979), p. 372.

If you follow the line of reasoning I used above, you may think that in order to make life easy for Susan, the mother should have used (226) instead of (227) in order to convey the intended information:

(226) He must think that your mother is rich.

(227) He must think I am rich.

The first formulation is naturally interpreted *de dicto*, and, on such a reading, it does not involve any violation of interpretation constraints. Thus, it is a more cooperative formulation than (227), since the intended interpretation of the latter involves a violation of *SHIFT, *ACCOMMODATION and STRENGTH in the described situation (see the OT-table relative to this example in the previous section). Still Susan's mother chooses to utter (227) and her choice is not unmotivated. First of all, Susan's mother is personally involved and probably upset because of the raise in tuition. It is *her* who has to pay more money now and the use of the personal pronoun 'I' rather than the neutral description 'your mother' is a more effective way to express her personal commitment and feelings about the situation. Furthermore, we might also assume that the following generation principle plays a role here, Zeevat (1999b) calls it **ParsePerson**, which expresses a preference for the use of the personal pronouns 'I', 'you' instead of descriptions in order to refer to the participants to the conversation. Principle **ParsePerson** might be related to a more general principle which expresses a preference for shorter utterances which require less effort to be generated. An utterance of (226) would violate such a principle whereas an utterance of (227) would not. From these considerations, it seems fair to conclude that in contrast to the previous presidential case where cooperative formulations were strictly preferred by the speaker, here, given the particular circumstances, none of the two formulations is strictly preferred over the other from Susan's mother perspective: the use of the description 'your mother' is more cooperative, but the use of the personal pronoun 'I' (i) is more effective for her in this particular situation in order to express her feelings and (ii) is shorter and thus possibly preferred because requires less effort to be generated.

This informal discussion suggests the following formalization of Susan's interpretation problem in such a situation:

$$\begin{array}{r|ccc|}
 & \mathit{UP(art)} & & \mathit{UP(moth)} \\ \hline
\text{'}\mathit{I}\text{'} & \bot & \rightarrow & \circ \\
 & \uparrow & & \\
\text{'}\mathit{your\ mother}\text{'} & \bot & \rightarrow & \circ \\ \hline
\end{array}$$

where

> $UP(art)$ is the result of updating the common ground with *de re* reading of (227) under the prominent cover containing the concept 'such and such a famous artist'.

> $UP(moth)$ is the result of updating the common ground with *de re* reading of (227) under a cover containing the concept 'Susan's mother'.

Note that $UP(moth)$ is equivalent to an update of the common ground with the *de dicto* reading of (226).

$UP(moth)$ is preferred over $UP(art)$ by Susan because the latter would lead to inconsistency. The use of the description '*your mother*', although it is more cooperative, is not strictly preferred over the use of the personal pronoun '*I*'' by Susan's mother because of the reasons discussed above.

Susan's mother utterance and her intended interpretation are Nash-optimal in such a game, as well as the profile containing the alternative cooperative formulation. The other two profiles are not even BJ-optimal.

4.4 Conclusion

In the first part of this chapter, I have looked more specifically at a number of formal properties of conceptual covers and I have compared them with alternative notions of cross-world identification. In the second part, I have attempted a first description of the pragmatics of conceptual covers. The proposed analysis is certainly not conclusive. A number of open questions remain in connection to the OT analysis discussed in the first part, for instance, concerning the choice of the constraints and their ranking. As for the game theoretical part, the field is new and most of the theoretical questions are still unsettled. For instance, already the characterization of the basic ingredients of an interpretation game is open to discussion. My identification of the set of possible actions for the addressee A_H with the set of possible update effects on the current common ground is not without consequences and is in need of further justification. Another open question is how generation constraints, interpretation principles and particular goals combine in the determination of the preference relations. All these issues are challenging and ask for further investigation, which, however, must be left to another occasion.

Appendix A

Proofs

A.1 Questions

A.1.1. FACT. [Rigidity and Answerhood] Let $M = \langle D, W \rangle$ be a standard model.

(i) $Pt \rhd_M ?xPx \quad \Leftrightarrow \quad t$ is rigid in M

(ii) $!Pt \ {}_c{\rhd}_M \ ?xPx \quad \Leftrightarrow \quad t$ is rigid in M

proof:

$\Leftarrow$ I give first an intuitive idea of why this direction holds. Suppose t is rigid in M. It then follows:

(i) the proposition associated with Pt in M, namely $\{w \mid [t]_{M,w} \in w(P)\}$ corresponds to the union of a non-empty subset of the blocks in the partition determined by $?xPx$, namely $\{w \mid d \in w(P)\}$ where $d = [t]_{M,w}$ for all $w \in W$. Thus Pt constitutes a partial answer to $?xPx$.

(ii) the proposition associated with $!Pt$ in M, $\{w \in W \mid w(P) = \{[t]_{M,w}\}\}$ corresponds to exactly one block in the partition determined by $?xPx$, namely $\{w \in W \mid w(P) = \{d\}\}$ where $d = [t]_{M,w}$ for all $w \in W$. Thus $!Pt$ constitutes a complete answer to $?xPx$.

I give now a more detailed proof of (i) $\Leftarrow$.

Suppose t is rigid in M and let $X = \{\alpha \in [\![?xPx]\!]_M \mid \alpha \subseteq [\![Pt]\!]_M\}$. Clearly:

(a) $X \neq \emptyset$, since, e.g., $\alpha = \{w \mid w(P) = \{[t]_{M,w}\}\} \in X$.

(b) $X \subset [\![?xPx]\!]_M$, since, $X \subseteq [\![?xPx]\!]_M$ by construction and, e.g., $\{w \mid w(P) = \emptyset\} \notin X$ and $\{w \mid w(P) = \emptyset\} \in [\![?xPx]\!]_M$.

(c) $[\![Pt]\!]_M = \cup X$. Indeed,

   (i) $\cup X \subseteq [\![Pt]\!]_M$ holds trivially by definition of X;

   (ii) $[\![Pt]\!]_M \subseteq \cup X$ holds because suppose $w \in [\![Pt]\!]_M$. Since $[\![?xPx]\!]_M$ is a partition of W, there must be a unique $\alpha \in [\![?xPx]\!]_M$ such that $w \in \alpha$. For any $w' \in \alpha$, $w(P) = w'(P)$, since $\alpha \in [\![?xPx]\!]_M$, and, since t is rigid, $w' \in [\![Pt]\!]_M$ as well. Since this holds for any $w' \in \alpha$, then $\alpha \in X$ and, therefore, $w \in \cup X$.

From (a), (b) and (c) we have that

$$\exists X \subset [\![?xPx]\!]_M : \ [\![Pt]\!]_M = \cup X \neq \emptyset$$

and therefore $Pt \rhd_M ?xPx$.

$\Rightarrow$ Suppose now t is not rigid in M. Let α be the set $\{w \in W \mid w(P) = \{d\}\}$ for some $d \in D$. α is obviously an element of $[\![?xPx]\!]_M$. Consider now the two worlds w and w' such that $w(P) = \{d\}$ and $w(t) = d$, and $w'(P) = \{d\}$ and $w'(t) = d'$ where $d \neq d'$. Since M is standard, and t is not rigid, w and w' will be elements of W and therefore, given the denotation they assign to P, they are elements of α. Obviously $w \in [\![Pt]\!]_M$ and $w \in [\![!Pt]\!]_M$, but $w' \notin [\![Pt]\!]_M$ and $w' \notin [\![!Pt]\!]_M$. This means that the following holds:

$$\exists \alpha \in [\![?xPx]\!]_M : (\alpha \bigcap [\![(!)Pt]\!]_M) \neq \emptyset \ \& \ \alpha \not\subseteq [\![(!)Pt]\!]_M$$

which implies that:

(i) $[\![Pt]\!]_M$ can not be equivalent to the union of a set of blocks in $[\![?xPx]\!]_M$, i.e. for no $X \subseteq [\![?xPx]\!]_M$ the following holds: $[\![Pt]\!]_M = \cup\{X\}$. Thus Pt does not partially answer $?xPx$, $Pt \not\rhd_M ?xPx$.

(ii) $[\![Pt]\!]_M$ is not an element of $[\![?xPx]\!]_M$, thus it does not completely answer the question, $!Pt \ {}_c\!\!\not\rhd_M ?xPx$. □

A.1.2. Fact. [Rigidity and Triviality]

$$t \text{ is rigid in } M \quad \Leftrightarrow \quad ?x\ t = x \text{ is trivial in } M$$

proof: This fact holds trivially, $?x\ t = x$ places those worlds in different blocks in which t denotes different individuals; however, if t is rigid and, only in this case, there are no such pairs of worlds, thus the question groups all worlds together.

A.1.3. FACT. [Cardinality] Let M be a model, $\wp$ be a conceptual perspective, g an assignment function and $\alpha \in [\![?\vec{x}\phi]\!]^{\wp}_{M}$ be a block in the partition determined by $?\vec{x}\phi$ in M under $\wp$. Then

$$\forall w, w' \in \alpha : |\lambda\vec{d}\,[\![\phi]\!]^{\wp}_{M,w,g[\vec{x}/\vec{d}]}| = |\lambda\vec{d}\,[\![\phi]\!]^{\wp}_{M,w',g[\vec{x}/\vec{d}]}|$$

proof: Let $f^{\vec{CC}}_{w,w'}$ be a relation between sequences $\vec{d}$ of n individuals figuring in w to sequences $\vec{d'}$ of n individuals figuring in w' such that

$$f^{\vec{CC}}_{w,w'}(\vec{d},\vec{d'}) \;\Leftrightarrow\; \exists\vec{c} \in \prod_{i\in n}(CC_i) :\; \vec{c}(w) = \vec{d}\;\&\;\vec{c}(w') = \vec{d'}$$

We prove that $f^{\vec{CC}}_{w,w'}$ is a total function and is bijective.

(i) Suppose $f^{\vec{CC}}_{w,w'}(\vec{d},\vec{d'})$ and $f^{\vec{CC}}_{w,w'}(\vec{d''},\vec{d'''})$. We show that

$$\vec{d} = \vec{d''} \;\Leftrightarrow\; \vec{d'} = \vec{d'''}$$

By definition of $f^{\vec{CC}}_{w,w'}$ we have that

$$\exists\vec{c},\vec{c'} \in \prod_{i\in n}(CC_i) : \vec{c}(w) = \vec{d}\;\&\;\vec{c}(w') = \vec{d'}\;\&\;\vec{c'}(w) = \vec{d''}\;\&\;\vec{c'}(w') = \vec{d'''}$$

$\Rightarrow$ Suppose $d = d''$. By the uniqueness condition on conceptual covers since $\vec{c}(w) = \vec{c'}(w)$ it follows that $\vec{c} = \vec{c'}$ which implies that $\vec{c}(w') = \vec{d'}$ is the same sequence as $\vec{c'}(w') = \vec{d'''}$.

$\Leftarrow$ As above.

We have proved that $f^{\vec{CC}}_{w,w'}$ is a function ($\Rightarrow$) and is injective ($\Leftarrow$).

(ii) By the existence condition on conceptual covers, $\forall\vec{d} \in D^n : \forall w \in W : \exists\vec{c} \in \prod_{i\in n}(CC_i) : \vec{c}(w) = \vec{d}$. Since the concepts in $\vec{c}$ are total functions, it holds that $\forall w' \in W : \exists\vec{d'} \in D^n : \vec{c}(w') = \vec{d'}$. Thus the following holds for all w and w':

$$\forall\vec{d} \in D^n : \exists\vec{c} \in \prod_{i\in n}(CC_i) : \exists\vec{d'} \in D^n : \vec{c}(w) = \vec{d}\;\&\;\vec{c}(w') = \vec{d'}$$

Therefore by definition of $f^{\vec{CC}}_{w,w'}$ it holds that

(a) $\forall\vec{d} \in D^n : \exists\vec{d'} \in D^n : f^{\vec{CC}}_{w,w'}(\vec{d},\vec{d'})$;

(b) $\forall\vec{d'} \in D^n : \exists\vec{d} \in D^n : f^{\vec{CC}}_{w,w'}(\vec{d},\vec{d'})$.

We have proved that $f^{\vec{CC}}_{w,w'}$ is total ($dom(f^{\vec{CC}}_{w,w'}) = D^n$ from (a)), and surjective ($range(f^{\vec{CC}}_{w,w'}) = D^n$ from (b)).

I will write $f^{\vec{CC}}_{w,w'}(\vec{d}) = \vec{d'}$ for $f^{\vec{CC}}_{w,w'}(\vec{d}, \vec{d'})$. Now, since $f^{\vec{CC}}_{w,w'}$ is a bijection, in order to prove proposition A.1.3, it is enough to prove the following lemma:

A.1.4. Lemma. Let $M = \langle D, W \rangle$ be a model, $\wp$ be a conceptual perspective, g an assignment function and $\alpha \in [\![?\vec{x}\phi]\!]^{\wp}_{M}$ be a block in the partition determined by $?\vec{x}\phi$ in M under $\wp$. Then

$$\forall w, w' \in \alpha : \forall \vec{d} \in D^n : [\![\phi]\!]^{\wp}_{M,w,g[\vec{x}/\vec{d}]} = 1 \quad \Leftrightarrow \quad [\![\phi]\!]^{\wp}_{M,w',g[\vec{x}/f^{\vec{CC}}_{w,w'}(\vec{d})]} = 1$$

proof: By definition of $f^{\wp(\vec{x})}_{w,w'}$, the following holds:

$$\forall \vec{d} \in D^n : \exists \vec{c} \in \prod_{i \in n}(\wp(\vec{x_i})) : \ \vec{c}(w) = \vec{d} \ \& \ \vec{c}(w') = f^{\wp(\vec{x})}_{w,w'}(\vec{d})$$

and this sequence $\vec{c}$ is clearly unique. Therefore $[\![\phi]\!]^{\wp}_{M,w,g[\vec{x}/\vec{d}]} = 1$ holds iff (i) holds:

(i) $[\![\phi]\!]^{\wp}_{M,w,g[\vec{x}/\vec{c}(w)]} = 1$

Since w and w' belong to the same block α in the partition determined by $?\vec{x}\phi$, by the semantics of interrogatives, (i) is the case iff the following is the case:

(ii) $[\![\phi]\!]^{\wp}_{M,w',g[\vec{x}/\vec{c}(w')]} = 1$

and (ii) clearly holds iff $[\![\phi]\!]^{\wp}_{M,w',g[\vec{x}/f^{\vec{CC}}_{w,w'}(\vec{d})]} = 1$. □

A.1.5. Fact. [K-Answerhood and Rigidity] Let $M = \langle D, W \rangle$ be a standard model.

$$Pt \ {}_{K}\!\rhd_{M} ?xPx \quad \Leftrightarrow \quad t \text{ is rigid in } M$$

proof: The proof follows trivially from the Karttunen semantics of constituent questions which, as is easily seen, assigns the following set of propositions as denotation to our question:

$${}_{K}[\![?xPx]\!]_{M,w,g} = \{\{v \mid d \in v(P)\} \mid d \in D \ \& \ d \in w(P)\}$$

$\Leftarrow$ Suppose t is rigid in M. This means that $[\![Pt]\!]_M = \{v \mid d \in v(P)\}$ for $d = [t]_M$. By the Karttunen analysis of constituent questions, this means $[\![Pt]\!]_M \in {}_{K}[\![?xPx]\!]_{M,w,g}$ for all $w \in W$ and g such that $[\![Pt]\!]_{M,w,g} = 1$. Since M is standard (and Pt is consistent) there is such a world w (and assignment g). Thus $\exists w \in W$,

$\exists g$: $[\![Pt]\!]_M \in {}_K[\![?xPx]\!]_{M,w,g}$, which means by the definition of singular positive answers $Pt\ {}_K\rhd_M ?xPx$.

$\Rightarrow$ Suppose t is not rigid in M. Then since M is standard, there is no $d \in D$ such that $[\![Pt]\!]_M = \{v \mid d \in v(P)\}$. So for no world w (and no assignment g), $[\![Pt]\!]_M$ is in the K-denotation of $?xPx$ in w (with respect to g), so the sentence does not constitute a singular positive answer to the question. $\square$

A.2 Belief

Comparison with MPL

A.2.1. Proposition. Let ϕ be a wff in $\mathcal{L}_{CC}$.

$$\models_{CCC} \phi \quad \text{iff} \quad \models_{MPL} \phi$$

proof: To prove this proposition I first show that given a classical CC-model M, we can define an equivalent ordinary modal predicate logic model M', that is, an MPL-model that satisfies the same wffs as M. Let M be $\langle W, R, D, I, \{CC\}\rangle$. We define an equivalent model $M' = \langle W', R', D', I'\rangle$ as follows. $W' = W$, $R' = R$, $D' = CC$. For I' we proceed as follows.

(i) $\forall \langle c_1, ..., c_n\rangle \in CC^n$, $w \in W$, $P \in \mathcal{P}$:

$$\langle c_1, ..., c_n\rangle \in I'(P)(w) \quad \text{iff} \quad \langle c_1(w), ..., c_n(w)\rangle \in I(P)(w)$$

(ii) $\forall c \in CC$, $w \in W$, $a \in \mathcal{C}$:

$$I'(a)(w) = c \quad \text{iff} \quad I(a)(w) = c(w)$$

Clause (ii) is well-defined because the uniqueness condition on covers guarantees that there is a unique $c \in CC$ such that $I(a)(w) = c(w)$.

In our construction we take the elements of the conceptual cover in the old model to be the individuals in the new model, and we stipulate that they do, in all w, what their instantiations in w do in the old model. Clause (i) says that a sequence of individuals is in the denotation of a relation P in w in the new model iff the sequence of their instantiations in w is in P in w in the old model. In order for clause (ii) to be well-defined, it is essential that CC is a conceptual cover, rather than an arbitrary set of concepts. In M', an individual constant a will denote in w the unique c in CC such that $I(a)(w) = c(w)$. That there is such a unique c is guaranteed by the uniqueness condition on conceptual covers. We have to prove that this construction works. I will use g, g' for assignments within M and h, h' for assignments within M'. Note that for all assignments g within M: $g(n) = CC$ for all CC-indices n, since CC is the unique cover available in M. I will say that g corresponds with h iff $g = h \cup \{\langle n, CC\rangle \mid n \in N\}$. This means that the two assignments assign the same values to all individual variables x_n for all n, and g assigns the cover CC to all CC-indices n.

A.2.2. THEOREM. Let g and h be any corresponding assignments. Let w be any world in W and ϕ any wff in $\mathcal{L}_{CC}$. Then

$$M, w, g \models_{CC} \phi \text{ iff } M', w, h \models_{MPL} \phi$$

proof: The proof is by induction on the construction of ϕ. We start by showing that the following holds for all terms t:

(A) $[t]_{M,w,g} = [t]_{M',w,h}(w)$

Suppose t is a variable. Then $[t]_{M,w,g} = g(t)(w)$. By definition of corresponding assignments, $g(t)(w) = h(t)(w)$. Since $[t]_{M',w,h} = h(t)$, this means that $[t]_{M,w,g} = [t]_{M',w,h}(w)$. Suppose now t is a constant. Then $[t]_{M,w,g} = I(t)(w)$. By the existence and uniqueness conditions on conceptual covers, there is a unique $c \in CC$, such that $I(t)(w) = c(w)$. By clause (ii) of the definition of I': $I'(t)(w) = c$. Since $[t]_{M',w,h} = I'(t)(w)$, this means $[t]_{M,w,g} = c(w) = [t]_{M',w,h}(w)$. We can now prove the lemma for atomic formulae.

Suppose ϕ is $Pt_1, ..., t_n$. Now $M, w, g \models_{CC} Pt_1, ..., t_n$ holds iff (a) holds:

(a) $\langle [t_1]_{M,w,g}, ..., [t_n]_{M,w,g}\rangle \in I(P)(w)$

By (A), (a) holds iff (b) holds:

(b) $\langle [t_1]_{M',w,h}(w), ..., [t_n]_{M',w,h}(w)\rangle \in I(P)(w)$

which, by definition of I', is the case iff (c) holds:

(c) $\langle [t_1]_{M',w,h}, ..., [t_n]_{M',w,h}\rangle \in I'(P)(w)$

which means that $M', w, h \models_{MPL} Pt_1, ..., t_n$.

Suppose now ϕ is $t_1 = t_2$. $M, w, g \models_{CC} t_1 = t_2$ holds iff (d) holds:

(d) $[t_1]_{M,w,g} = [t_2]_{M,w,g}$

By (A) above, (d) holds iff (e) holds:

(e) $[t_1]_{M',w,h}(w) = [t_2]_{M',w,h}(w)$

which, by the uniqueness condition on conceptual covers, is the case iff (f) holds:

(f) $[t_1]_{M',w,h} = [t_2]_{M',w,h}$

which means that $M', w, h \models_{MPL} t_1 = t_2$.

The lemma is proved for atomic cases. The induction for $\neg$, $\exists$, $\wedge$ and $\Box$ is immediate. $\Box$

Now it is clear that if a classical CC-model M and an ordinary MPL-model M' correspond in the way described then the theorem entails that any wff in $\mathcal{L}_{CC}$ is CC-valid in M iff it is MPL-valid in M'. Thus, given a classical CC-model, we can define an equivalent MPL-model, but also given an MPL-model, we can define an equivalent classical CC-model $\langle W, R, D, \{CC\}, I\rangle$ by taking CC to be the rigid cover. This suffices to prove proposition A.2.1. $\Box$

Axiomatization

Recall our definition of a CC-theorem. CC consists of the following axiom schemata:

Basic propositional modal system

PC All propositional tautologies.

K $\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$

Quantifiers Recall that $\phi[t]$ and $\phi[t']$ differ only in that the former contains the term t in one or more places where the latter contains t'.

EGa $\phi[t] \to \exists x_n \phi[x_n]$ (if ϕ is atomic)

EGn $\phi[y_n] \to \exists x_n \phi[x_n]$

BFn $\forall x_n \Box\phi \to \Box\forall x_n \phi$

Identity

ID $t = t$

SIa $t = t' \to (\phi[t] \to \phi[t'])$ (if ϕ is atomic)

SIn $x_n = y_n \to (\phi[x_n] \to \phi[y_n])$

LNIn $x_n \neq y_n \to \Box\, x_n \neq y_n$

Let AX_{CC} be the set of axioms of CC. The set of CC-theorems T_{CC} is the smallest set such that:

AX $AX_{CC} \subseteq T_{CC}$

MP If ϕ and $\phi \to \psi \in T_{CC}$, then $\psi \in T_{CC}$

∃I If $\phi \to \psi \in T_{CC}$ and x_n not free in ψ, then $(\exists x_n \phi) \to \psi \in T_{CC}$

N If $\phi \in T_{CC}$, then $\Box\phi \in T_{CC}$

I use the standard notation and write $\vdash_{CC} \phi$ for $\phi \in T_{CC}$.

I list now a number of theorems and rules that will be used later on, I omit the derivations which are standard.

First of all, by **∃I** and **EGn**, we can derive all standard predicate logic theorems and rules governing the behaviour of quantifiers which do not involve any shift of index, among others the rule of introduction of the universal quantifier **∀I**, and the principle of renaming of bound variables **PRn**, and the two principles **PL1n** and **PL2n**.

∀I $\vdash_{CC} \phi \to \psi \;\; \Rightarrow \; \vdash_{CC} \phi \to \forall x_n \psi$ (provided x_n is not free in ϕ)

PRn $\vdash_{CC} \exists x_n \phi \leftrightarrow \exists y_n \phi[x_n/y_n]$

PL1n $\vdash_{CC} \forall x_n(\phi \to \psi) \to (\phi \to \forall x_n \psi)$ (provided x_n not free in ϕ)

PL2n $\vdash_{CC} \exists z_n(\exists x_n \phi \to \phi[x_n/z_n])$ (provided z_n not free in $\exists x_n \phi$)

Furthermore, we can prove the following two identity theorems. The derivation of **ID'** uses **ID** and **SIa**, whereas the derivation of **LIn** uses **ID**, **N** and **SIn**.

LIn $\vdash_{CC} x_n = y_n \to \Box\, x_n = y_n$

ID' $\vdash_{CC} t = t' \to t' = t$

Recall our definition of CC-validity.

A.2.3. Definition. [CC-Validity] Let ϕ be in $\mathcal{L}_{CC}$. Then

$$\models_{CC} \phi \;\text{ iff }\; M \models_{CC} \phi \quad \text{for all CC-models } M$$

Soundness

A.2.4. Theorem. [Soundness] If $\vdash_{CC} \phi$, then $\models_{CC} \phi$.

proof: The proof that **MP**, **∃I** and **N** preserve validity is standard. The validity of **PC** and **K** is obvious. The validity of **EGa** and **SIa** follows from proposition 2.4.9. **LNIn** is valid by proposition 2.4.8. The validity of **SIn** may be established by induction on the construction of ϕ. We show that **EGn** and **BFn** are valid.

EGn Suppose $M, w \models_g \phi[y_n]$. This implies $M, w \models_{g[x_n/g(y_n)]} \phi[x_n]$. Since $g(y_n) \in g(n)$, $M, w \models_g \exists x_n \phi[x_n]$.

BFn Suppose $M, w \models_g \forall x_n \Box\phi$. Let h be any x_n-alternative of g, $g[x_n]h$, i.e. let h be such that: $\forall v \in (N \cup \mathcal{V}_N)$: $v \neq x_n \;\Rightarrow\; g(v) = h(v)$ and $h(x_n) \in h(n)$. Let wRw'. Then $M, w \models_h \Box\phi$, and hence $M, w' \models_h \phi$. Since this holds for every x_n-alternative h of g, we have $M, w' \models_g \forall x_n \phi$; and since this holds for all w' such that wRw', we finally have $M, w \models_g \Box\forall x_n \phi$. □

Completeness

We show that for any ϕ which is not a theorem in CC we can define a CC-model in which ϕ is not valid. The technique we will use in the construction of these models varies only slightly from the standard technique used for modal predicate logic with identity (see in particular Hughes and Cresswell (1996)). The worlds of these models will be maximal CC-consistent sets of wffs which have the *witness property*, that is, for all indices n, for all wffs of the form $\exists x_n \phi$, there is a n-indexed variable y_n such that $\exists x_n \phi \to \phi[x_n/y_n] \in w$. In order to obtain this we will follow the common practice and consider an expanded language $\mathcal{L}^+$, which is $\mathcal{L}_{CC}$ with the addition of a denumerable set of fresh variables.

We assume the standard results about maximal consistent sets of wffs with respect to a system S.

A.2.5. Theorem. [Lindenbaum's Theorem] Any S-consistent set of wffs Δ can be enlarged to a maximal S-consistent set of wffs Γ.

A.2.6. Theorem. Suppose Γ is a maximal consistent set of wffs with respect to S. Then

1. for each wff ϕ, exactly one member of $\{\phi, \neg\phi\}$ belongs to Γ.

2. for each pair of wffs ϕ and ψ, $\phi \wedge \psi \in \Gamma$ iff $\phi \in \Gamma$ and $\psi \in \Gamma$.

3. if $\vdash_S \phi$, then $\phi \in \Gamma$.

4. if $\phi \in \Gamma$ and $\vdash_S \phi \to \psi$, then $\psi \in \Gamma$.

The basic theorem about the witness property is the following:

A.2.7. THEOREM. Let Λ be a consistent set of wffs of $\mathcal{L}_{CC}$. Let $\mathcal{V}^+$ be a denumerable set of new variable symbols and let $\mathcal{L}^+$ be the simple expansion of $\mathcal{L}_{CC}$ formed by adding $\mathcal{V}^+$. Then there is a consistent set Δ of wffs of $\mathcal{L}^+$ with the witness property such that $\Lambda \subseteq \Delta$.

proof: I follow the standard proof (see Hughes and Cresswell (1996), pp. 258–259), the only difference is that we have not just one sort of variables, but many. We assume that all wffs of the form $\exists x_n \phi$ for any wff ϕ of $\mathcal{L}^+$ and any index n, and any variable x_n for any index n are enumerated so that we can speak of the first, the second and so on. We define a sequence of sets Δ_0, Δ_1, ... etc. as follows:

$$\Delta_0 = \Lambda$$

$$\Delta_{m+1} = \Delta_m \cup \{\exists x_n \phi \to \phi[x_n/y_n]\}$$

where $\exists x_n \phi$ is the $m+1$th wff in the enumeration of wff of that form and y_n is the first variable indexed with n not in Δ_m or in ϕ. Since Δ_0 is in $\mathcal{L}$ and Δ_m has been formed from it by the addition of only m wffs there will be infinitely many n-indexed variables from ${\mathcal{V}^+}_n$ to provide such a y_n.

Δ_0 is assumed to be consistent so we shall show that Δ_{m+1} is if Δ_m is. Suppose not. Then there will be $\psi_1 \wedge ... \wedge \psi_m$ in Δ_m such that:

(i) $\vdash_{CC} (\psi_1 \wedge ... \wedge \psi_m) \to \exists x_n \phi$

(ii) $\vdash_{CC} (\psi_1 \wedge ... \wedge \psi_m) \to \neg\phi[x_n/y_n]$

Since y_n does not occur in Δ_m, it is not free in $(\psi_1 \wedge ... \wedge \psi_m)$ and so from (ii) by $\forall$**I** we have:

(iii) $\vdash_{CC} (\psi_1 \wedge ... \wedge \psi_m) \to \forall y_n \neg\phi[x_n/y_n]$

which can be rewritten as:

(iv) $\vdash_{CC} (\psi_1 \wedge ... \wedge \psi_m) \to \neg\exists y_n \phi[x_n/y_n]$

Now since y_n did not occur in ϕ, $\exists y_n \phi[x_n/y_n]$ is a bound alphabetic variant of $\exists x_n \phi$, and so by **PRn**:

(v) $\vdash_{CC} (\psi_1 \wedge ... \wedge \psi_m) \to \neg\exists x_n \phi$

But (i) and (v) give

(vi) $\vdash_{CC} \neg(\psi_1 \wedge ... \wedge \psi_m)$

which contradicts the consistency of Δ_m. Let Δ be the union of all the Δ_ms. It is easy to see that Δ is consistent and has the witness property. $\Box$

Once a set Δ has the witness property each extension of Δ in the same language also has the witness property. Lindenbaum's theorem guarantees that if Δ is consistent there is a maximal consistent set Γ such that $\Delta \subseteq \Gamma$, and so since Δ has the witness property, Γ does too. The standard result we can prove about maximal consistent sets with the witness property in modal logic is the following:

A.2.8. THEOREM. If Γ is a maximal consistent set of wffs in some language (say $\mathcal{L}^+$) of modal predicate logic, and Γ has the witness property, and α is a wff such that $\Box\alpha \notin \Gamma$, then there is a consistent set Δ of wff of $\mathcal{L}^+$ with the witness property such that $\{\psi \mid \Box\psi \in \Gamma\} \bigcup \{\neg\alpha\} \subseteq \Delta$.

proof: Again we can use the standard construction (see Hughes and Cresswell (1996), pp. 259–261). Again we assume that all wffs of the form $\exists x_n\phi$ for any wff ϕ of $\mathcal{L}^+$ and any index n, and any variable x_n for any index n are enumerated so that we can speak of the first, the second and so on. We define a sequence of wffs γ_0, γ_1, γ_2, ... etc. as follows:

γ_0 is $\neg\alpha$

γ_{m+1} is $\gamma_m \wedge (\exists x_n\phi \rightarrow \phi[x_n/y_n])$

where $\exists x_n\phi$ is the $m+1$th wff in the enumeration of that form and y_n is the first variable indexed by n such that:

(*) $\{\psi \mid \Box\psi \in \Gamma\} \cup \{\gamma_m \wedge (\exists x_n\phi \rightarrow \phi[x_n/y_n])\}$ is consistent.

In order for this construction to succeed we have to be sure that there always will be a n-indexed variable y_n satisfying (*).

Since γ_0 is $\neg\alpha$, $\{\psi \mid \Box\psi \in \Gamma\} \cup \{\gamma_0\}$ is consistent from a standard result of propositional modal logic. We show that provided $\{\psi \mid \Box\psi \in \Gamma\} \cup \{\gamma_m\}$ is consistent, there always will be a n-indexed variable y_n satisfying (*).

Suppose there were not. Then for every variable y_n in $\mathcal{V}_n^+$, there will exist some $\{\psi_1, ..., \psi_k\} \subseteq \{\psi \mid \Box\psi \in \Gamma\}$ such that

$$\vdash_{CC} (\psi_1 \wedge ... \wedge \psi_k) \rightarrow (\gamma_m \rightarrow \neg(\exists x_n\phi \rightarrow \phi[x_n/y_n]))$$

so, by propositional modal logic (**N**, **K** and $\Box$-distribution):

$$\vdash_{CC} (\Box\psi_1 \wedge ... \wedge \Box\psi_k) \rightarrow \Box(\gamma_m \rightarrow \neg(\exists x_n\phi \rightarrow \phi[x_n/y_n]))$$

But Γ is maximal consistent and $\Box\psi_1, ..., \Box\psi_k \in \Gamma$, and so

(i) $\Box(\gamma_m \rightarrow \neg(\exists x_n\phi \rightarrow \phi[x_n/y_n])) \in \Gamma$

Let z_n be some n-indexed variable not occurring in ϕ or in γ_m. Since Γ has the witness property, then we have for some n-witness y_n:

$$\exists z_n(\neg\Box(\gamma_m \rightarrow \neg(\exists x_n\phi \rightarrow \phi[x_n/z_n]))) \rightarrow \neg\Box(\gamma_m \rightarrow \neg(\exists x_n\phi \rightarrow \phi[x_n/y_n])) \in \Gamma$$

Since (i) holds for all $y_n \in \mathcal{V}_n^+$, we then have that:

$$\neg\exists z_n\neg\Box(\gamma_m \rightarrow \neg(\exists x_n\phi \rightarrow \phi[x_n/z_n])) \in \Gamma$$

But Γ is CC-maximal consistent and hence by **BFn** we have:

$$\Box\forall z_n(\gamma_m \rightarrow \neg(\exists x_n\phi \rightarrow \phi[x_n/z_n])) \in \Gamma$$

Since z_n does not occur γ_m, then by **PL1n** we have:

$$\Box(\gamma_m \rightarrow \neg\exists z_n(\exists x_n\phi \rightarrow \phi[x_n/z_n])) \in \Gamma$$

But since z_n does not occur in ϕ by **PL2n**, we have

$$\vdash_{CC} \exists z_n(\exists x_n\phi \rightarrow \phi[x_n/z_n])$$

but then $\Box\neg\gamma_m \in \Gamma$ and so $\neg\gamma_m \in \{\psi \mid \Box\psi \in \Gamma\}$ which would make $\{\psi \mid \Box\psi \in \Gamma\} \cup \{\gamma_m\}$ inconsistent against our assumption.

Let Δ be the union of $\{\psi \mid \Box\psi \in \Gamma\}$ and all the γ_ms. Since each $\{\psi \mid \Box\psi \in \Gamma\} \cup \{\gamma_m\}$ is consistent, and since $\vdash_{CC} \gamma_m \rightarrow \gamma_k$ for $m \geq k$, so is their union Δ. Any maximal consistent extension of Δ has all the required properties and so the theorem is proved. $\Box$

I will now show that for each CC-consistent set Δ of wffs of $\mathcal{L}$, we can construct a model M_Δ containing a world in which all the wffs in Δ are true. These models M_Δ are based on *cohesive* sub-frames of the frame $F = \langle W_F, R_F\rangle$ where:

(a) W_F is the set of CC-maximal consistent sets of wffs of $\mathcal{L}^+$ which have the witness property.

(b) wR_Fw' iff for every wff $\Box\phi$ of $\mathcal{L}^+$, if $\Box\phi \in w$, then $\phi \in w'$.

Cohesive frames are frame in which each two worlds are linked by means of some forwards or backwards R-chain. The reason why we need to consider cohesive models is that in a cohesive model for CC for each index n any world verifies exactly the same identity formulas between variables in $\mathcal{V}_n$.

Let Δ be a CC-consistent set of wffs of $\mathcal{L}_{\mathcal{CC}}$. We show how to construct $M_\Delta = \langle W, R, D, I, C\rangle$ in which there is a world w^* such that $\Delta \subseteq w^*$. The construction of W, R, D, I are standard. C will be some extra work.

W Given some world $w^* \in W_F$ such that $\Delta \subseteq w^*$, we let W be the set of all and only those worlds in W_F which are reachable from w^* by a chain of forwards R_F-steps. These worlds are maximal consistent sets of wffs of $\mathcal{L}^+$, which satisfy the witness property.

R R is R_F restricted to W.

D Let $\sim$ be the following relation over the set $\mathcal{V}_0^+$ of variables of $\mathcal{L}^+$ with index 0.

$$v_0 \sim x_0 \text{ iff } v_0 = x_0 \in w$$

Since W is cohesive (every w, w' in W are linked by some R-chain) and every w contains **LIn** and **LNIn**, it makes no difference which w is selected for this purpose. We can prove the following lemma:

A.2.9. Lemma. $\forall w, w' \in W : \forall n \in N : \forall x, y \in \mathcal{V}^+ : \ x_n = y_n \in w$ iff $x_n = y_n \in w'$

proof: Consider any two worlds w, w' such that wRw' or $w'Rw$. Suppose $x_n = y_n \in w$. If wRw', then by **LIn**, $\Box x_n = y_n \in w$ and so $x_n = y_n \in w'$. If $w'Rw$, then if $x_n = y_n \notin w'$, then $x_n \neq y_n \in w'$ and so by **LNIn** $\Box x_n \neq y_n \in w'$ and so $x_n = y_n \notin w$, contradicting the assumption. Now since our M_Δ is cohesive, then any two worlds w and w' in W are linked by a chain of backwards or forwards R-steps, and so if $x_n = y_n \in w$ then $x_n = y_n \in w'$, and if $x_n = y_n \in w'$ then $x_n = y_n \in w$. $\Box$

It is easy to see that $\sim$ is an equivalence relation (we use **ID** and **SIa** and the maximal consistency of each w). Now for each $x_0 \in \mathcal{V}_0^+$, let

$$[x_0] = \{y_0 \in \mathcal{V}_0^+ \mid x_0 \sim y_0\}$$

be the equivalence class of x_0. We take the domain D to be the set of all these equivalence classes $[x_0]$, for $x_0 \in \mathcal{V}_0^+$ and so define $D = \{[x_0] \mid x_0 \in \mathcal{V}_0^+\}$.

I We now define the interpretation function I for the predicate and individual constant symbols of the language.

(i) For each n-placed relation symbol P in $\mathcal{L}_{CC}$, for each $w \in W$ we define the interpretation $I(P)(w)$ of the symbol P in w as follows: $\forall \langle [x_{0_1}], ..., [x_{0_n}] \rangle \in D^n$,

$$\langle [x_{0_1}], ..., [x_{0_n}] \rangle \in I(P)(w) \text{ iff } Px_{0_1}, ..., x_{0_n} \in w$$

The definition is independent of the representatives of the equivalence classes $[x_{0_1}], ..., [x_{0_n}]$, because by **SIa** (or **SIn**) we have

$$\vdash_{CC} Px_{0_1}, ..., x_{0_n} \wedge x_{0_1} = y_{0_1} \wedge ... \wedge x_{0_n} = y_{0_n} \rightarrow Py_{0_1}, ..., y_{0_n}$$

(ii) Let a be a constant symbol of $\mathcal{L}$ and $w \in W$. From **ID** and **EGa**, we have

$$\vdash_{CC} \exists x_0(a = x_0)$$

So $\exists x_0(a = x_0) \in w$, and because w has the witness property, there is a 0-indexed variable y_0 such that

$$a = y_0 \in w$$

y_0 may not be unique, but its equivalence class $[y_0]$ is unique because, using **SIa** we have:

$$\vdash_{CC} a = y_0 \wedge a = z_0 \rightarrow y_0 = z_0$$

The interpretation $I(a)(w)$ in w is this (uniquely determined) element $[y_0]$ of D.

$$I(a)(w) = [y_0] \quad \text{iff} \quad a = y_0 \in w$$

C Let c_{v_n} be the function from W to D such that for all $w \in W$:

$$c_{v_n}(w) = [y_0] \quad \text{iff} \quad v_n = y_0 \in w$$

The proof that for all w there is such a unique element $[y_0] \in D$ is parallel to the one in the second clause of the definition of I. By **ID** and **EGa**, we have $\vdash_{CC} \exists x_0(v_n = x_0)$ and so $\exists x_0(v_n = x_0) \in w$ for all w, and, therefore, by the witness property of w, $v_n = y_0 \in w$ for some 0-witness y_0. y_0 may not be unique, but its equivalence class $[y_0]$ is, because, by using **SIa**, we have: $\models_{CC} v_n = y_0 \wedge v_n = z_0 \rightarrow z_0 = y_0$.

We let now $CC_n = \{c_{v_n} \mid v_n \in \mathcal{V}_n^+\}$ and we define $C = \{CC_n \mid n \in N\}$.

We have to show that these sets CC_n are conceptual covers.

(i) Existence Condition: $\forall w \in W : \forall\, [x_0] \in D : \exists c_{v_n} \in CC_n : c_{v_n}(w) = [x_0]$.
proof: Take any w and $[x_0]$. By **ID** and **EGa**, we have $\exists y_n(y_n = x_0) \in w$ and because w has the witness property we know that $v_n = x_0 \in w$ for some n-witness v_n. Consider now c_{v_n} which is in CC_n. By definition $c_{v_n}(w) = [x_0]$.

(ii) Uniqueness Condition: $\forall w \in W : \forall c_{v_n} c_{z_n} \in CC_n : c_{v_n}(w) = c_{z_n}(w) \Rightarrow c_{v_n} = c_{z_n}$
proof: Note firstly that for all $w' \in W$ the following holds:
(A) $c_{x_n}(w') = c_{y_n}(w') \Leftrightarrow x_n = y_n \in w'$

($\Rightarrow$) by definition of c_{x_n} and c_{y_n}, and **SIa**; ($\Leftarrow$) suppose $x_n = y_n \in w'$ by **ID**, **EGa**, **SIa**, and witness property of w', we have for some z_0 and v_0, $x_n = z_0 \in w'$, $y_n = v_0 \in w'$ and $z_0 = v_0 \in w'$, which by definition of c_{x_n} and c_{y_n} means $c_{x_n}(w') = c_{y_n}(w')$.
Suppose now $c_{v_n}(w) = c_{z_n}(w)$ for some $c_{v_n}, c_{z_n} \in CC_n$, and w. By (A) this implies $v_n = z_n \in w$, which, by lemma A.2.9, implies that for any $w' \in W$, $v_n = z_n \in w'$ and so, again by (A), $c_{v_n}(w') = c_{z_n}(w')$. Since this holds for all $w' \in W$, we have $c_{v_n} = c_{z_n}$.

We define the *canonical assignment* g as follows: $\forall n \in N, \forall x_n \in \mathcal{V}_n^+$: $g(n) = CC_n$ and $g(x_n) = c_{x_n}$. We can now prove the following theorem:

A.2.10. Theorem. For any $w \in W$, and any wff $\phi \in \mathcal{L}^+$,

$$M_\Delta, w \models_g \phi \quad \text{iff} \quad \phi \in w$$

proof: the proof is by induction on the construction of ϕ. I start by showing that (B) holds for all t in $\mathcal{L}^+$:

(B) $[t]_{M_\Delta,w,g} = [x_0]$ iff $t = x_0 \in w$

Suppose t is an indexed variable v_n in $\mathcal{V}_N^+$. Then $[t]_{M_\Delta,w,g} = g(v_n)(w)$. By definition of canonical assignment $g(v_n)(w) = c_{v_n}(w)$. By definition of c_{v_n}, $c_{v_n}(w) = [x_0]$ iff $v_n = x_0 \in w$.

Suppose t is a constant symbol a in $\mathcal{L}$. Then $[t]_{M_\Delta,w,g} = I(a)(w)$. By clause (ii) in the definition of I, $I(a)(w) = [x_0]$ iff $a = x_0 \in w$. We can now prove the theorem for atomic formulae.

(a) Consider $Rt_1, ..., t_n$. Let $w \in W$. By (B), the denotation of the terms $t_1, ..., t_n$ in w will be some $[x_{0_1}], ..., [x_{0_n}]$ where $t_1 = x_{0_1}, ..., t_n = x_{0_n} \in w$. Then

$$M_\Delta, w \models_g Rt_1, ..., t_n \Leftrightarrow \langle [x_{0_1}], ..., [x_{0_n}] \rangle \in I(R)(w) \Leftrightarrow Rx_{0_1}, ..., x_{0_n} \in w$$

But $t_1 = x_{0_1}, ..., t_n = x_{0_n} \in w$, thus by various application of **SIa** we have that

$$Rx_{0_1}, ..., x_{0_n} \leftrightarrow Rt_1, ..., t_n \in w$$

and so $Rx_{0_1}, ..., x_{0_n} \in w \Leftrightarrow Rt_1, ..., t_n \in w$.

(b) $M_\Delta, w \models_g t_1 = t_2$ iff $[t_1]_{M_\Delta,w,g} = [t_1]_{M_\Delta,w,g}$. By (B) above this is the case iff $[x_{0_1}] = [x_{0_2}]$ for some $[x_{0_1}]$ and $[x_{0_2}]$ such that $t_1 = x_{0_1} \in w$ and $t_2 = x_{0_2} \in w$. Obviously $x_{0_1} = x_{0_2} \in w$, and therefore by various applications of **SIa** we have that $t_1 = t_2 \in w$.

(c) $M_\Delta, w \models_g \neg\phi$ iff $M_\Delta, w \not\models_g \phi$ iff $\phi \notin w$ iff $\neg\phi \in w$.

(d) $M_\Delta, w \models_g \phi \wedge \psi$ iff $M_\Delta, w \models_g \phi$ and $M_\Delta, w \models_g \psi$ iff $\phi \in w$ and $\psi \in w$ iff $\phi \wedge \psi \in w$.

(e) Suppose $\exists x_n \phi \in w$. By the witness property, for some y_n, we have $\phi[x_n/y_n] \in w$. But then by induction hypothesis $M_\Delta, w \models_g \phi[x_n/y_n]$ which, by standard principle of replacement, implies $M_\Delta, w \models_{g[x_n/g(y_n)]} \phi$. Since $g(y_n) = c_{y_n}$ is an element of $g(n)$ this implies $M_\Delta, w \models_g \exists x_n \phi$.

Suppose $M_\Delta, w \models_g \exists x_n \phi$. Then $M_\Delta, w \models_{g[x_n/c_{v_n}]} \phi$ for some $c_{v_n} \in g(n)$. Since by definition of canonical assignment $g(v_n) = c_{v_n}$, by standard principle of replacement, we have $M_\Delta, w \models_g \phi[x_n/v_n]$. But then, by induction hypothesis, $\phi[x_n/v_n] \in w$ and by **EGn** $\exists x_n \phi \in w$.

(f) Suppose $\Box\phi \in w$ and wRw'. Then $\phi \in w'$ and so $M_\Delta, w' \models_g \phi$. Since this holds for all w' such that wRw', we have $M_\Delta, w \models_g \Box\phi$.

Suppose $\Box\phi \notin w$. Then $\neg\Box\phi \in w$. But then by theorem A.2.8 (in combination with A.2.5) we know that there is some $w' \in W_F$ such that wR_Fw' and $\phi \notin w'$. w' is clearly in W as well since it is accessible from w. Thus by induction hypothesis $M_\Delta, w' \not\models_g \phi$. Since wRw', we can conclude $M_\Delta, w' \not\models_g \Box\phi$. □

A.2.11. Theorem. [Completeness] If $\models_{CC} \phi$, then $\vdash_{CC} \phi$

proof: Suppose $\not\vdash_{CC} \phi$. Then $\neg\phi$ is CC-consistent. We then know that $\neg\phi$ is an element of some world of the model $M_{\{\neg\phi\}}$ generated by $\{\neg\phi\}$ and therefore by theorem A.2.10 true in some world in that model. This means that $M_{\neg\phi} \not\models_{CC} \phi$. Since $M_{\neg\phi}$ is a CC-model, $\not\models_{CC} \phi$. □

We have shown that the system CC is sound and complete with respect to the class of all CC-models. By standard techniques we can show that CC+**D**+**4**+**E** is sound and complete with respect to all serial, transitive and euclidean CC-models.

A.3 Dynamics

A.3.1. Proposition. Let ϕ be a novel sentence. Then

$$\approx\!\!\!\!\models_{old} \phi \;\;\Leftrightarrow\;\; \approx\!\!\!\!\models_{new} \phi$$

proof: One direction of the proof hinges on the fact that (α) given a new state s, a new assignment g and a novel sentence ϕ, we can construct an old state σ and perspective $\wp$, such that if $s \approx\!\!\!\!\models_g \phi$ then $\sigma \approx\!\!\!\!\models_\wp \phi$. For the other direction, we show that (β) given an old state σ connected under a perspective $\wp$ we can find a new state s and a new assignment g such that for all ϕ, if $\sigma \models\!\!\!\approx_{\wp} \phi$, then $s \models\!\!\!\approx_{g} \phi$. The two constructions are straightforward.

(α) Let s be a new state, g be a new assignment, and $X \subseteq \mathcal{V}_N$ a set of indexed variables. Let $a_{(g,w,X)}$ be an old assignment such that $dom(a_{(g,w,X)}) = X$ and $\forall v \in X : a_{(g,w,X)}(v) = g(v)(w)$. Then $\sigma_{(s,g,X)}$ is the following set: $\{\langle w, a_{(g,w,X)}\rangle \mid w \in s\}$; and $\wp_{(g)}$ is such that for all indices $n \in N$: $\wp_{(g)}(n) = g(n)$.

(β) Suppose now σ is an old state CC-accessible under a perspective $\wp$. Then let $s_\sigma = \{w \in W \mid \exists a : \langle w, a\rangle \in \sigma\}$. Suppose v is an indexed variable defined in σ. I will write $\sigma(v)$ to denote the function $f : s_\sigma \to D$ such that $\forall w \in s_\sigma$: $f(w) = a(v)$ where a is the unique assignment such that $\langle w, a\rangle \in \sigma$. The uniqueness of such an a is guaranteed by the definiteness of σ, since σ is an accessible state. Now let $G_{(\sigma,\wp)}$ be the following set $\{g \in (C^N \cup (D^W)^{\mathcal{V}_N}) \mid \forall n \in N : g(n) = \wp(n)\ \&\ \forall v \in dom(\sigma) : \forall w \in s_\sigma :\ g(v)(w) = \sigma(v)(w)\}$. The $\wp$-uniformity of σ guarantees that for all n and x_n, $g(x_n) \in g(n)$. Thus any g in $G_{(\sigma,\wp)}$ is a new assignment.

We can prove the following theorem: (I write $QV(\phi)$ to denote the set of quantified variables in ϕ, $AQV(\phi)$ to denote the set of dynamically active quantified variables in ϕ and $FV(\phi)$ to denote the set of variables in ϕ which are neither syntactically nor dynamically bound in ϕ)

A.3.2. THEOREM. Let ϕ be novel.

(α) Let $X \subseteq \mathcal{V}_N$ be such that $FV(\phi) \subseteq X$ and $QV(\phi) \cap X = \emptyset$. Then

$$s[\phi]^g_h = t \quad\Rightarrow\quad \sigma_{(s,g,X)}[\phi]^{\wp(g)}\sigma_{(t,h,(X\cup AQV(\phi)))}$$

(β) Let σ be CC-accessible under $\wp$ and $g \in G_{(\sigma,\wp)}$. Then

$$\sigma[\phi]^{\wp}\tau \quad\Rightarrow\quad \exists h \in G_{(\tau,\wp)} : s_\sigma[\phi]^g_h = s_\tau$$

proof: the proof is by mutual induction on the complexity of ϕ. I will just give the atomic cases **(a)** $Rt_1, ..., t_n$, and **(b)** $\exists x_n$ and the induction step for **(c)** dynamic conjunction. Notice that the existential quantifier can be treated as an atomic update also in the old formulation of the semantics:

$$\sigma[\exists x_n]^{\wp}\tau \quad \text{iff} \quad \sigma[x_n/c]\tau \text{ for some } c \in \wp(n)$$

In definition 3.4.2, I chose the classical format for defining existential quantification in uniformity with the definitions in 3.2.8 of the other styles of dynamic quantification, in which the existential was not treated as an atomic action, since MS quantification could not have been formulated in such a fashion.

(a) ϕ is $Rt_1, ...t_n$.

(α) Suppose $s[Rt_1, ...t_n]^g_h = t$. This is the case iff the following holds:

$$g = h \ \ \& \ \ t = \{w \in s \mid \langle [t_1]_{w,g}, ..., [t_n]_{w,g}\rangle \in w(R)\}$$

By construction of $\sigma_{(t,h,X)}$, this means:

$$\text{(i)} \ \ \sigma_{(t,h,X)} = \{\langle w, a_{(g,w,X)}\rangle \mid w \in s \ \& \ \langle [t_1]_{w,g}, ..., [t_n]_{w,g}\rangle \in w(R)\}$$

Notice that for all terms t defined in $\langle w, a_{(g,w,X}\rangle$ the following holds:

$$\text{(ii)} \ \ [t]_{w,g} = \langle w, a_{(g,w,X)}\rangle(t)$$

Indeed, if t is a constant: $[t]_{w,g} = w(t) = \langle w, a_{(g,w,X))}\rangle(t)$. If t is a variable, and if t is in X, then $[t]_{w,g} = g(t)(w) = a_{(g,w,X)}(t) = \langle w, a_{(g,w,X))}\rangle(t)$.
Since, by assumption, $FV(Rt_1, ..., t_n) \subseteq X$, from (i) and (ii), we then have:

$$\sigma_{(t,h,X)} = \{\langle w, a_{(g,w,X)}\rangle \mid w \in s \ \& \ \langle \langle w, a_{(g,w,X)}\rangle(t_1), ..., \langle w, a_{(g,w,X)}\rangle(t_n)\rangle \in w(R)\}$$

By construction of $\sigma_{(s,g,X)}$, this implies:

$$\sigma_{(t,h,X)} = \{i \in \sigma_{(s,g,X)} \mid \langle i(t_1), ..., i(t_n)\rangle \in i(R)\}$$

which means for all $\wp$:

$$\sigma_{(s,g,X)}[Rt_1, ..., t_n]^{\wp}\sigma_{(t,h,X)}$$

So in particular for $\wp_{(g)}$. Since $(X \cup AQV(Rt_1, ...t_n)) = X$, we then have:

$$\sigma_{(s,g,X)}[Rt_1, ..., t_n]^{\wp(g)}\sigma_{(t,h,(X\cup AQV(Rt_1,...,t_n)))}$$

(β) Suppose $\sigma[Rt_1, ..., t_n]^{\wp}\tau$. Then $\tau = \{i \in \sigma \mid \langle i(t_1), ..., i(t_n)\rangle \in i(R)\}$. By definition of s_τ, this means:

$$s_\tau = \{w \in W \mid \exists a : \langle w, a\rangle \in \sigma \ \& \ \langle \langle w, a\rangle(t_1), ..., \langle w, a\rangle(t_n)\rangle \in w(R)\}$$

By definition of s_σ and, since $g \in G_{(\sigma,\wp)}$, we have $\langle w, a\rangle = [t]_{w,g}$. So we have:

$$s_\tau = \{w \in s_\sigma \mid \langle [t_1]_{w,g}, ..., [t_n]_{w,g}\rangle \in w(R)\}$$

which means:

$$s_\sigma[Rt_1, ..., t_n]^g_g = s_\tau$$

Since $\tau \subseteq \sigma$, then $G_{(\sigma,\wp)} \subseteq G_{(\tau,\wp)}$. Thus $\exists h \in G_{(\tau,\wp)} : s_\sigma[Rt_1, ..., t_n]^g_h = s_\tau$.

(b) ϕ is $\exists x_n$.

(α) Suppose $s[\exists x_n]^g_h = t$. Then $g[x_n]h$ and $t = s$. By definition of $g[x_n]h$ and construction of $\sigma_{(t,h,(X\cup\{x_n\}))}$, we then have (note that by assumption X does not contain x_n):

$$\sigma_{(t,h,(X\cup\{x_n\}))} = \{\langle w, a\rangle \mid w \in s \ \& \ dom(a) = X \cup \{x_n\} \ \& \ a(x_n) = h(x_n)(w) \ \& \ \forall v \in X : a(v) = g(v)(w)\}$$

But this means:

$$\sigma_{(t,h,(X\cup\{x_n\}))} = \{\langle w, a_{(g,w,X)}\rangle[x_n/c(w)] \mid w \in s \ \& \ c = h(x_n)\}$$

which implies by construction of $\sigma_{(s,g,X)}$ and since $h(n) = g(n)$ for all n:

$$\sigma_{(t,h,(X\cup\{x_n\}))} = \{i[x_n/c(w_i)] \mid i \in \sigma_{(s,g,X)} \ \& \ c \in g(n)\}$$

which implies by definition of c-extension and of $\wp_{(g)}$:

$$\sigma_{(s,g,X)}[x_n/c]\sigma_{(t,h,(X\cup\{x_n\}))} \text{ for some } c \in \wp_{(g)}$$

which means $\sigma_{(s,g,X)}[\exists x_n]^{\wp(g)}\sigma_{(t,h,(X\cup AQV(\exists x_n)))}$.

(β) Suppose $\sigma[\exists x_n]^{\wp}\tau$. Then $\sigma[x_n/c]\tau$ for some $c \in \wp(n)$, which means the following:

(i) $\tau = \{i[x_n/c(w_i)] \mid i \in \sigma \ \& \ c \in \wp(n)\}$

Since no world occurring in σ is eliminated in τ, a consequence of (i) is (ii):

(ii) $s_\sigma = s_\tau$

Consider now $g \in G_{(\wp,\sigma)}$. It is easy to see that:

(iii) any assignment h such that $g[x_n]h$ and $\forall w \in s_\tau : \ h(x_n)(w) = \tau(x_n)(w)$ is in $G_{(\wp,\tau)}$.

Clause (iii) follows from the fact that given (i), $G_{(\wp,\tau)}$ is the following set:

$$G_{(\wp,\tau)} = \{h \in (C^N \cup (D^W)^{\mathcal{V}_N}) \mid \forall n \in N : h(n) = \wp(n) \ \& \ \forall w \in s_\tau : \ \forall v \in dom(\tau) : ((v \neq x_n \ \Rightarrow h(v)(w) = \sigma(v)(w)) \ \& \ (v = x_n \ \Rightarrow \ h(v)(w) = \tau(v)(w)))\}$$

from (ii) and (iii), it follows that $s_\sigma = s_\tau$ and $\exists h \in G_{(\tau,\wp)} : g[x_n]h$, which means $\exists h \in G_{(\tau,\wp)} : s_\sigma[\exists x_n]^g_h = s_\tau$.

(c) ϕ is $\psi_1 \wedge \psi_2$.

(α) Suppose now $s[\psi_1 \wedge \psi_2]^g_h = t$. This means $\exists k : (s[\psi_1]^g_k)[\psi_2]^k_h = t$. By induction hypothesis we then have:

$$\exists k : \sigma_{(s,g,Y)}[\psi_1]^{\wp(g)}\sigma_{((s[\psi_1]^g_k),k,(Y\cup AQV(\psi_1)))} \;\&\; \sigma_{((s[\psi_1]^g_k),k,Z)}[\psi_2]^{\wp(k)}\sigma_{(t,h,(Z\cup AQV(\psi_2)))}$$

where

(i) $FV(\psi_1) \subseteq Y$ & $QV(\psi_1) \cap Y = \emptyset$ and

(ii) $FV(\psi_2) \subseteq Z$ & $QV(\psi_2) \cap Z = \emptyset$

Consider any $X \subseteq \mathcal{V}_N$ such that $FV(\psi_1 \wedge \psi_2) \subseteq X$ and $QV(\psi_1 \wedge \psi_2) \cap X = \emptyset$. It is easy to see that X satisfies (i) and $(X \cup AQV(\psi_1))$ satisfies (ii) (recall that $\psi_1 \wedge \psi_2$ is novel). But then we have for all such X:

$$\exists k : \sigma_{(s,g,X)}[\psi_1]^{\wp(g)}\sigma_{((s[\psi_1]^g_k),k,(X\cup AQV(\psi_1)))} \;\&\; \sigma_{((s[\psi_1]^g_k),k,(X\cup AQV(\psi_1)))}[\psi_2]^{\wp(k)}\sigma_{(t,h,((X\cup AQV(\psi_1))\cup AQV(\psi_2)))}$$

Since $\wp_{(g)} = \wp_{(k)}$ this implies:

$$\exists \tau : \sigma_{(s,g,X)}[\psi_1]^{\wp(g)}\tau \;\&\; \tau[\psi_2]^{\wp(g)}\sigma_{(t,h,(X\cup AQV(\psi_1\wedge\psi_2))}$$

which means $\sigma_{(s,g,X)}[\psi_1 \wedge \psi_2]^{\wp(g)}\sigma_{(t,h,(X\cup AQV(\psi_1\wedge\psi_2))}$.

(β) Suppose $\sigma[\psi_1 \wedge \psi_2]^{\wp}\tau$. Then $\exists\chi$: $\sigma[\psi_1]^{\wp}\chi[\psi_2]^{\wp}\tau$. By induction hypothesis, we have the following two facts:

(i) $\exists k \in G_{(\chi,\wp)} : s_\sigma[\psi_1]^g_k = s_\chi$

and

(ii) $\forall k' \in G_{(\chi,\wp)} : \exists h \in G_{(\tau,\wp)} : s_\chi[\psi_2]^{k'}_h = s_\tau$

which imply:

(iii) $\exists h \in G_{(\tau,\wp)} : \exists k : (s_\sigma[\psi_1]^g_k)[\psi_2]^k_h = s_\tau$

which means $\exists h \in G_{(\tau,\wp)} : s_\sigma[\psi_1 \wedge \psi_2]^g_h = s_\tau$. The other induction steps are left to the reader. Notice that case of negation requires the mutual induction. $\Box$

A.3.3. Corollary. Let ϕ be novel and $X \subseteq \mathcal{V}_N$ be such that $FV(\phi) \subseteq X$ and $QV(\phi) \cap X = \emptyset$. Then

$$s \approx\!\!\!\models_g \phi \;\; \Rightarrow \;\; \sigma_{(s,g,X)} \approx\!\!\!\models_{\wp(g)} \phi$$

proof: Suppose $s \approx\!\!\!\models_g \phi$. Then $\exists h : s[\phi]^g_h = s$. Since ϕ is novel and X is specified as above, by theorem A.3.2, clause (α), this means $\exists h : \sigma_{(s,g,X)}[\phi]^{\wp(g)}\sigma_{(s,h,(X \cup AQV(\phi)))}$. Since for each possibility $\langle w, a_{(g,w,X)}\rangle \in \sigma_{(s,g,X)}$, it holds that

(i) $\langle w, a_{(h,w,(X \cup AQV(\phi)))}\rangle \in \sigma_{(s,h,(X \cup AQV(\phi)))}$ (by construction of the two states)

(ii) $a_{(g,w,X)} \subseteq a_{(h,w,(X \cup AQV(\phi)))}$ (since g and h may differ only in the values they assign to the variables in $AQV(\phi)$)

we then have that $\sigma_{(s,g,X)} \prec \sigma_{(s,h,(X \cup AQV(\phi)))}$. It follows that $\sigma(s, g, \phi) \approx\!\!\!\models_{\wp(g)} \phi$. □

A.3.4. Corollary. Let σ be CC-accessible under $\wp$ and $g \in G_{(\wp,\sigma)}$. Then

$$\sigma \approx\!\!\!\models_{\wp} \phi \;\; \Rightarrow \;\; s_\sigma \approx\!\!\!\models_g \phi$$

proof: Suppose $\sigma \approx\!\!\!\models_{\wp} \phi$. Then $\exists \tau$: $\sigma[\phi]_\wp \tau$ & $\sigma \prec \tau$. Since σ is CC-accessible and $g \in G_{(\wp,\sigma)}$, this implies, by theorem A.3.2, clause (β), that $\exists h$: $s_\sigma[\phi]^g_h = s_\tau$. Since $\sigma \prec \tau$, we have $s_\sigma = s_\tau$. It follows that $s_\sigma \approx\!\!\!\models_g \phi$. □

The proof of proposition A.3.1 is now a trivial exercise. Suppose $\approx\!\!\!\models_{new} \phi$ and ϕ is novel. Then $\exists M$,g, $\exists s \in S_M$: $s \neq \emptyset$ & $s \approx\!\!\!\models_g \phi$. Since ϕ is novel, by corollary A.3.3, it follows that for all $X \subseteq \mathcal{V}_N$ such that $FV(\phi) \subseteq X$ and $QV(\phi) \cap X = \emptyset$: $\sigma_{(s,g,X)} \approx\!\!\!\models_{\wp(g)} \phi$. Furthermore, since $s \neq \emptyset$, then such a $\sigma_{(s,g,X)} \neq \emptyset$, which means $\approx\!\!\!\models_{old} \phi$. Suppose now $\approx\!\!\!\models_{old} \phi$. Then $\exists M$, $\exists \wp, \exists \sigma \in \Sigma_M$: $\mathrm{Acc}_\wp(\sigma)$ & $\sigma \neq \emptyset$ & $\sigma \approx\!\!\!\models_\wp \phi$. By corollary A.3.4, it follows that $\forall g \in G_{(\wp,\sigma)}$: $s_\sigma \approx\!\!\!\models_g \phi$. Since $G_{(\wp,\sigma)} \neq \emptyset$, and, since $\sigma \neq \emptyset$, also $s_\sigma \neq \emptyset$, we then have $\approx\!\!\!\models_{new} \phi$. □

A.4 Formal and Pragmatic Aspects of Conceptual Covers

A.4.1. Proposition. Let CC be a cover over (W, D). The method of cross-identification R_{CC} determined by CC is proper.

proof: Given a conceptual cover CC over $\langle W, D\rangle$ the corresponding R_{CC} is defined as follows:

$$\langle w, d\rangle R_{CC}\langle w', d'\rangle \text{ iff } \exists c \in CC : c(w) = d \,\&\, c(w') = d'$$

We have to prove that R_{CC} satisfies conditions (i) and (ii) on proper methods of cross-identification:

(i) R_{CC} is an equivalence relation:

   (a) Reflexivity: $\forall w, d : \exists c \in CC : c(w) = d$ (by existence) $\Rightarrow \forall w, d : \langle w, d\rangle R_{CC}\langle w, d\rangle$;

   (b) Symmetry: $\langle w, d\rangle R_{CC}\langle w', d'\rangle \Rightarrow$ (by construction) $\exists c \in CC : c(w) = d \,\&\, c(w') = d' \Rightarrow \langle w', d'\rangle R_{CC}\langle w, d\rangle$;

   (c) Transitivity: $\langle w, d\rangle R_{CC}\langle w', d'\rangle$ and $\langle w', d'\rangle R_{CC}\langle w'', d''\rangle \Rightarrow \exists c : c(w) = d$ and $c(w') = d'$ and $\exists c' : c'(w') = d'$ and $c'(w'') = d'' \Rightarrow$ (by uniqueness) $c = c' \Rightarrow \exists c : c(w) = d$ and $c(w'') = d'' \Rightarrow \langle w, d\rangle R_{CC}\langle w'', d''\rangle$.

(ii) Each individual has one and only counterpart in each world:

   (a) $\forall w, w', d : \exists d' : \langle w, d\rangle R\langle w', d'\rangle$:
   By existence $\forall w, d : \exists c \in CC : c(w) = d$ and since c is a total function $\exists d' : c(w') = d' \Rightarrow \forall w, d : \forall w' : \exists d' : \langle w, d\rangle R_{CC}\langle w', d'\rangle$;

   (b) $\forall w, w', d, d', d'' : \langle w, d\rangle R\langle w', d'\rangle \,\&\, \langle w, d\rangle R\langle w', d''\rangle \;\Rightarrow\; d' = d''$:
   $\langle w, d\rangle R_{CC}\langle w', d'\rangle \,\&\, \langle w, d\rangle R_{CC}\langle w', d''\rangle \Rightarrow$ (by construction) $\exists c : c(w) = d \,\&\, c(w') = d'$, and $\exists c' : c'(w) = d \,\&\, c'(w') = d'' \Rightarrow$ (by uniqueness) $c = c' \Rightarrow d' = d''$. □

A.4.2. PROPOSITION. The set of classes of pairs CP_R induced by a proper cross-identification method R is a conceptual cover.

proof: The set of classes of pairs induced by R is the following set:

$$CP_R = \{[w, d]_R \mid w \in W \,\&\, d \in D\}$$

where $[w, d]_R = \{\langle w', d'\rangle \mid \langle w, d\rangle R\langle w', d'\rangle\}$.
The result in A.4.2 follows from the following two lemmas:

A.4.3. LEMMA. Let R be a proper method of cross-identification over (W, D). Then $\forall \alpha \in CP_R$, $\forall w \in W$, $\exists! d \in D$: $\langle w, d\rangle \in \alpha$.

proof: For any α, by construction, $\alpha = [w, d]_R$ for some w, d, and by condition (iia), there is a $\langle w, d\rangle$ in α for all w. Suppose now $\langle w, d\rangle \in \alpha$ and $\langle w, d'\rangle \in \alpha$. This means by construction of CP_R that for some w'', d'', $\langle w'', d''\rangle R\langle w, d\rangle$ and $\langle w'', d''\rangle R\langle w, d'\rangle$. Since R is an equivalence relation: $\langle w, d\rangle R\langle w, d'\rangle$, which implies by condition (iib) that $d = d'$. □

A.4.4. LEMMA. Let R be a proper method of cross-identification over (W, D). Then $\forall w \in W$, $\forall d \in D$, $\exists! \alpha \in CP_R$: $\alpha(w) = d$.

proof: Since R is an equivalence relation, CP_R is a partition of the set of all world-individual pairs. This means that $\forall w, d : \exists! \alpha \in CP_R$: $\langle w, d \rangle \in \alpha$. $\square$

Bibliography

Aloni, M.: 1997a. Quantification in dynamic semantics. *in* P. Dekker, M. Stokhof and Y. Venema (eds), *Proceedings of the Eleventh Amsterdam Colloquium*. ILLC, University of Amsterdam. Amsterdam.

Aloni, M.: 1997b. Quantification in dynamic semantics. *in* E. Kraak and R. Wassermann (eds), *Proceedings of Accolade '97*. OzsL. Amsterdam. 12 page version.

Aloni, M.: 1998. A semantics for belief attributions. *Technical report*. University of Amsterdam. Contribution to Workshop on Communication and Attitudes.

Aloni, M.: 2000. Conceptual covers in dynamic semantics. *in* P. Blackburn and J. Seligman (eds), *Logic, Language and Computation, Vol. III*. CSLI. Stanford, CA.

Aloni, M.: to appear. Questions under cover. *in* D. Barker-Plummer, D. Beaver, J. van Benthem and P. S. de Luzio (eds), *Logic Unleashed: Language, Diagrams, and Computation*. CSLI, Stanford, CA.

Aloni, M., Beaver, D. and Clark, B.: 1999. Topic and focus sensitive operators. *in* P. Dekker (ed.), *Proceedings of the Twelfth Amsterdam Colloquium*. ILLC, University of Amsterdam. Amsterdam.

Barwise, J. and Perry, J.: 1983. *Situations and Attitudes*. MIT Press. Cambridge, MA.

Beaver, D.: 1994. When variables don't vary enough. *in* L. Santelmann and M. Harvey (eds), *Proceedings of the Fourth Semantics and Linguistic Theory Conference*. DMLL Publications, Cornell University.

Beaver, D.: 1995. *Presupposition and Assertion in Dynamic Semantics*. PhD thesis. University of Edinburgh. To appear: Studies in Logic, Language and Information, CSLI, Stanford, CA.

Beaver, D. I.: 1992. The kinematics of presupposition. *in* P. Dekker and M. Stokhof (eds), *Proceedings of the Eighth Amsterdam Colloquium*. ILLC, University of Amsterdam. Amsterdam.

Beaver, D. I.: 1993. Kinematic Montague grammar. *in* H. Kamp (ed.), *Presupposition*. ILLC, University of Amsterdam. Amsterdam. Dyana-2 deliverable R2.2.A, Part II.

Belnap, N. and Steel, T.: 1976. *The Logic of Questions and Answers*. Yale University Press. New Haven, CT.

Berman, S.: 1991. *On the Semantics and Logical Form of Wh-Clauses*. PhD thesis. University of Massachusetts at Amherst.

Blutner, R. 1998. Lexical pragmatics. *Journal of Semantics* **15**, 115–162.

Blutner, R.: 1999. Some aspects of optimality in natural language interpretation. *in* H. de Hoop and H. de Swart (eds), *OTS Papers on Optimality Theoretic Semantics*. OTS, Utrecht University. Utrecht. [to appear in *Journal of Semantics*].

Blutner, R. and Jäger, G.: 1999. Competition and interpretation: the german adverbs of repetition. *Extended conference Abstracts: Approaching the Grammar of Adjuncts*. University of Oslo.

Boër, S. and Lycan, W.: 1985. *Knowing Who*. MIT Press. Cambridge, MA.

Bonomi, A.: 1983. *Eventi Mentali*. Il Saggiatore. Milano.

Bonomi, A.: 1995. Transparency and specificity in intensional contexts. *in* P. Leonardi and M. Santambrogio (eds), *On Quine, New Essays*. Cambridge University Press. Cambridge, MA. pp. 164–185.

Carnap, R.: 1947. *Meaning and Necessity*. Chicago University Press. Chicago.

Chierchia, G. 1992. Anaphora and dynamic binding. *Linguistics and Philosophy* **15**, 111–183.

Chierchia, G.: 1995. *Dynamics of Meaning: anaphora, presupposition, and the theory of grammar*. University of Chicago Press. Chicago.

Church, A. 1982. A remark concerning Quine's paradox about modality. *Analisis Filosofico* **2**, 25–32. Spanish version.

Crimmins, M. and Perry, J. 1989. The prince and the phone booth: reporting puzzling beliefs. *Journal of Philosophy* **86**, 685–711.

de Hoop, H. and de Swart, H.: 1999. Temporal adjunct clauses in optimality theory. *in* H. de Hoop and H. de Swart (eds), *OTS Papers on Optimality Theoretic Semantics*. OTS, Utrecht University. Utrecht. [to appear in *Journal of Semantics*].

de Hoop, H. and Hendriks, P.: 1999. Optimality theoretic semantics. [to appear in *Linguistics and Philosophy*].

Dekker, P.: 1993. *Transsentential Meditations. Ups and Downs in Dynamic Semantics*. PhD thesis. University of Amsterdam.

Dekker, P.: 1994. Predicate logic with anaphora (seven inch version). *in* L. Santelmann and M. Harvey (eds), *Proceedings of the Fourth Semantics and Linguistic Theory Conference*. DMLL Publications, Cornell University. p. 17.

Dekker, P. 1996. The values of variables in dynamic semantics. *Linguistics and Philosophy* **19**(3), 211–257.

Dekker, P.: 1997. On first order information exchange. *in* A. Benz and G. Jäger (eds), *Proceedings of MunDial97*. University of Munich.

Dekker, P. 1998. Speaker's reference, descriptions and information structure. *Journal of Semantics* **15**(4), 301–330.

Dekker, P.: 2000a. Coreference and representationalism. *in* K. von Heusinger and U. Egli (eds), *Reference and Anaphorical Relations*. Kluwer. Dordrecht.

Dekker, P.: 2000b. Meaning and use of indefinite expressions. [to appear in *Journal of Logic, Language and Computation*].

Dekker, P. and van Rooy, R.: 1999. Bi-directional optimality theory: An application of game theory. *in* H. de Hoop and H. de Swart (eds), *OTS Papers on Optimality Theoretic Semantics*. OTS, Utrecht University. Utrecht. [to appear in *Journal of Semantics*].

Evans, G. 1977. Pronouns, quantifiers and relative clauses (1). *The Canadian Journal of Philosophy* **7**, 467–536. [Reprinted in: Evans, G., 1985, *Collected Papers*, Reidel, Dordrecht, pp. 76–152].

Fagin, R., Halpern, J. Y., Moses, Y. and Vardi, M.: 1995. *Reasoning about Knowledge*. MIT Press. Cambridge, MA.

Frege, G. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* **100**, 25–50.

Gerbrandy, J.: 1997. Questions of identity. *in* P. Dekker, M. Stokhof and Y. Venema (eds), *Proceedings of the Eleventh Amsterdam Colloquium.* ILLC, University of Amsterdam. Amsterdam.

Gerbrandy, J.: 1999. *Bisimulations on Planet Kripke.* PhD thesis. University of Amsterdam.

Gerbrandy, J.: 2000. Identity in epistemic semantics. *in* P. Blackburn and J. Seligman (eds), *Logic, Language and Computation, Vol. III.* CSLI. Stanford, CA.

Geurts, B. 1996. Local satisfaction guaranteed: a presupposition theory and its problems. *Linguistics and Philosophy* **19**, 259–294.

Ginzburg, J. 1995. Resolving questions, I & II. *Linguistics and Philosophy* **18**, 459–527, 567–609.

Ginzburg, J.: 1996. Interrogatives: Questions, facts and dialogues. *in* S. Lappin (ed.), *Handbook of Contemporary Semantic Theory.* Blackwell. Oxford. pp. 385–422.

Groenendijk, J.: 1998. Questions in update semantics. *in* J. Hulstijn and A. Nijholt (eds), *Formal Semantics and Pragmatics of Dialogue.* University of Twente.

Groenendijk, J.: 1999. The logic of interrogation. *in* T. Matthews and D. Strolovitch (eds), *The Proceedings of the Ninth Conference on Semantics and Linguistic Theory.* CLC Publications.

Groenendijk, J. and Stokhof, M.: 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers.* PhD thesis. University of Amsterdam.

Groenendijk, J. and Stokhof, M. 1991. Dynamic predicate logic. *Linguistics and Philosophy* **14**(1), 39–100.

Groenendijk, J. and Stokhof, M.: 1997. Questions. *in* J. van Benthem and A. ter Meulen (eds), *Handbook of Logic and Language.* Elsevier. Amsterdam.

Groenendijk, J., Stokhof, M. and Veltman, F.: 1996. Coreference and modality. *in* S. Lappin (ed.), *Handbook of Contemporary Semantic Theory.* Blackwell. Oxford.

Gupta, A.: 1980. *The Logic of Common Nouns: an investigation in quantified modal logic.* Yale University Press. New Haven, CT.

Hamblin, C. L. 1958. Questions. *Australasian Journal of Philosophy* **36**, 159–168.

Hamblin, C. L. 1973. Questions in Montague English. *Foundation of Language* **10**, 41–53.

Hausser, R. and Zaeffer, D.: 1979. Questions and answers in a context-dependent Montague grammar. *in* F. Guenthner and S. Schmidt (eds), *Formal Semantics and Pragmatics for Natural Languages*. Reidel. Dordrecht. pp. 339–358.

Heim, I.: 1982. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis. University of Massachusetts at Amherst. [Published in 1989 by Garland, New York].

Heim, I.: 1983a. File change semantics and the familiarity theory of definiteness. *in* R. Bäuerle, C. Schwarze and A. von Stechow (eds), *Meaning, Use, and Interpretation of Language*. de Gruyter. Berlin. pp. 164–189.

Heim, I.: 1983b. On the projection problem for presuppositions. *in* M. Barlow, D. Flickinger and M. Wescoat (eds), *Proceedings of WCCFL II*. Stanford, CA.

Heim, I. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* **9**, 183–221.

Heim, I.: 1994. Interrogative semantics and Karttunen's semantics for *know*. *in* R. Buchalla and A. Mittwoch (eds), *The Proceedings of the Ninth Annual Conference and the Worshop on Discourse of the Israel Association for Theoretical Linguistics*. Academon. Jerusalem.

Higginbotham, J.: 1991. Interrogatives I. *MIT Working Papers in Linguistics*. Vol. 15. MIT. Cambridge, MA.

Higginbotham, J.: 1996. Questions. *in* S. Lappin (ed.), *Handbook of Contemporary Semantic Theory*. Blackwell. Oxford.

Higginbotham, J. and May, R. 1981. Questions, quantifiers and crossing. *The Linguistic Review* **1**, 41–80.

Higgins, F. R.: 1973. *The Pseudo-Cleft Construction in English*. PhD thesis. MIT.

Hintikka, J.: 1962. *Knowledge and Belief*. Cornell UP. Ithaca, MA.

Hintikka, J.: 1967. On the logic of perception. *in* N. S. Care and R. M. Grimm (eds), *Perception and Personal Identity*. The Press of Case Western Reserve University. [Reprinted in: Hintikka J., 1969, *Models for Modalities*, Reidel, Dordrecht].

Hintikka, J.: 1969. Semantics for propositional attitudes. *in* Davis, Hockney and Wilson (eds), *Philosophical Logic*. Reidel. Dordrecht.

Hintikka, J.: 1975. *The Intensions of Intensionality and Other New Models for Modalities*. Reidel. Dordrecht.

Hintikka, J. 1976. The semantics of questions and the questions of semantics. *Acta Philosophica Fennica.*

Horn, L.: 1984. Towards a new taxonomy for pragmatic inference: Q-based and R-based implicatures. *in* D. Schiffrin (ed.), *Meaning, Form, and Use in Context.* Georgetown University Press. Washington. pp. 11–42.

Hughes, G. E. and Cresswell, M. J.: 1996. *A New Introduction to Modal Logic.* Routledge. London and New York.

Hulstijn, J.: 1997. Structured information states. Raising and resolving issues. *in* A. Benz and G. Jäger (eds), *Proceedings of MunDial97.* University of Munich.

Jäger, G.: 1995. Only updates. *in* P. Dekker and M. Stokhof (eds), *Proceedings of the Tenth Amsterdam Colloquium.* ILLC, University of Amsterdam. Amsterdam.

Janssen, T. M.: 1984. Individual concepts are useful. *in* F. Landman and F. Veltman (eds), *Varieties of Formal Semantics.* Foris. Dordrecht.

Kamp, H.: 1981. A theory of truth and semantic representation. *in* J. Groenendijk, T. Janssen and M. Stokhof (eds), *Formal Methods in the Study of Language.* Mathematical Centre. Amsterdam. [Reprinted in J. Groenendijk, T. Janssen, and M. Stokhof (eds.), *Truth, Interpretation and Information.* Foris, Dordrecht, 1984].

Kamp, H. and Reyle, U.: 1993. *From Discourse to Logic.* Kluwer. Dordrecht.

Kaplan, D.: 1969. Quantifying in. *in* D. Davidson and J. Hintikka (eds), *Words and Objections: Essays on the Work of W. V. Quine.* Reidel. Dordrecht. pp. 221–243.

Kaplan, D.: 1978. Dthat. *in* P. Cole (ed.), *Syntax and Semantics 9: Pragmatics.* Academic Press. New York. pp. 221–243.

Karttunen, L. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* **1**, 3–44.

Karttunen, L. and Peters, S.: 1976. Conventional implicatures in Montague grammar. *in* C. Oh and D. Dineen (eds), *Syntax and Semantics 11: Presupposition.* Academic Press. New York.

Kraut, R. 1983. There are no *de dicto* attitudes. *Synthese* **54**, 275–294.

Krifka, M.: 1999. For a structured account of questions and answers. *in* C. Smith (ed.), *Proceedings to Workshop on Spoken and Written Text.* University of Texas at Austin.

Kripke, S. 1963. Semantical considerations on modal logics. *Acta Philosophica Fennica* **16**, 83–94.

Kripke, S.: 1972. Naming and necessity. *in* D. Davidson and G. Harman (eds), *Semantics of Natural Languages*. Reidel. Dordrecht. chapter 254–355. Revised and enlarged revision first published in 1980 by Blackwell, Oxford.

Kripke, S.: 1979. A puzzle about belief. *in* A. Margalit (ed.), *Meaning and Use*. Reidel. Dordrecht. pp. 239–283.

Lahiri, U.: 1991. *Embedded Interrogatives and Predicates That Embed Them*. PhD thesis. MIT.

Lewis, D. 1968. Counterpart theory and quantified modal logic. *The Journal of Philosophy* **65**, 113–126.

Lewis, D. 1979. Attitudes de dicto and de se. *The Philosophical Review* **88**(4), 513–545.

Lewis, D. 1983. Individuation by acquaintance and by stipulation. *Philosophical Review* **92**, 3–12.

Neale, S.: 1993. *Descriptions*. MIT Press. Cambridge, MA.

Osborne, M. and Rubistein, A.: 1994. *A Course in Game Theory*. MIT Press. Cambridge, MA.

Prince, A. and Smolensky, P. 1997. Optimality: from neural network to universal grammar. *Science* **275**, 1604–1610.

Quine, W. V.: 1953. *From a Logical Point of View*. Harvard University Press. chapter VII Reference and modality.

Quine, W. V. 1956. Quantifiers and propositional attitudes. *Journal of Philosophy* **53**, 101–111. [Reprinted in: W.V. Quine, 1966, *The Ways of Paradox and Other Essays*, Random House, New York].

Quine, W. V.: 1960. *Word and Object*. Cambridge, MA.

Quine, W. V. 1961. Reply to Professor Marcus. *Synthese* **13**, 323–330. [Reprinted in: W.V. Quine, 1966, *The Ways of Paradox and Other Essays*, Random House, New York].

Richards, M. 1993. Direct reference and ascription of belief. *Journal of Philosophical Logic* **16**, 123–148.

Roberts, C.: 1996a. Anaphora in intensional contexts. *in* S. Lappin (ed.), *Handbook of Contemporary Semantic Theory*. Blackwell. Oxford.

Roberts, C.: 1996b. Information structure in discourse: Towards an integrated formal theory of pragmatics. *in* J. Yoon and A. Kathol (eds), *OSU Working Papers in Linguistics 49*. pp. 91–136.

Rooth, M. 1992. A Theory of Focus Interpretation. *Natural Language Semantics* **1**, 75–116.

Russell, B. 1905. On denoting. *Mind* **14**, 479–493.

Slater, B. 1986. E-type pronouns and ϵ-terms. *Canadian Journal of Philosophy* **16**, 27–38.

Stalnaker, R.: 1978. Assertion. *in* P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*. Academic Press. New York. pp. 315–332.

Stalnaker, R. 1981. Indexical belief. *Synthese* **49**, 129–151.

Stalnaker, R. 1987. Counterparts and identity. *Midwest Studies in Philosophy, Studies in Essentialism* **11**, 121–140.

Stalnaker, R.: 1988. Belief attribution and context. *in* R. Grimm and D. Merril (eds), *Contents of Thought*. Arizona University Press. Tucson.

Tichy, P. 1978. Questions, answers and logic. *American Philosophical Quarterly* **15**, 275– 284.

van Benthem, J., Muskens, R. and Visser, A.: 1997. Dynamics. *in* J. van Benthem and A. ter Meulen (eds), *Handbook of Logic and Language*. Elsevier Science Publishers.

van der Sandt, R. A. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* **9**, 333–377.

van Eijck, J. and Cepparello, G.: 1994. Dynamic modal predicate logic. *in* M. Kanazawa and C. Piñon (eds), *Dynamics, Polarity and Quantification*. CSLI. Stanford, CA.

van Fraassen, B. 1979. Propositional attitudes in weak pragmatics. *Journal of Philosophy* **76**, 365–374.

van Leeuwen, J.: 1991. *Individuals and Sortal Concepts*. PhD thesis. University of Amsterdam.

van Rooy, R.: 1997. *Attitudes and Changing Contexts*. PhD thesis. University of Stuttgart.

van Rooy, R.: 1999. Questioning to resolve decision problems. *in* P. Dekker (ed.), *Proceedings of the Twelfth Amsterdam Colloquium*. ILLC, University of Amsterdam. Amsterdam.

van Rooy, R.: 2000a. Decision problems in pragmatics. *in* M. Poesio and D. Traum (eds), *Proceedings of Götalog 2000*. Göteborg University. Göteborg. pp. 151–158.

van Rooy, R.: 2000b. The pragmatics of mention-some readings of questions. manuscript, Amsterdam.

Veltman, F. 1997. Defaults in update semantics. *Journal of Philosophical Logic* **25**, 221–261.

Vermeulen, C. F. 1996. Merging without mystery or: Variables in dynamic semantics. *Journal of Philosophical Logic* **24**(4), 405–450.

von Stechow, A.: 1990. Focusing and backgrounding operators. *in* W. Abraham (ed.), *Discourse Particles*. number 6. John Benjamins. Amsterdam. pp. 37–84.

von Stechow, A. and Zimmermann, E. 1984. Term answers and contextual changes. *Linguistics* **22**, 3–40.

Westerståhl, D.: 1984. Determiners and context sets. *in* J. Van Benthem and A. Ter Meulen (eds), *Generalized Quantifiers in Natural Language*. Foris. Dordrecht. pp. 45–71.

Williams, E. 1997. Blocking and anaphora. *Linguistic Inquiry* **28**, 577–628.

Zeevat, H.: 1995. A neoclassical analysis of belief sentences. *in* P. Dekker and M. Stokhof (eds), *Proceedings of the Tenth Amsterdam Colloquium*. ILLC, University of Amsterdam. Amsterdam.

Zeevat, H.: 1999a. The asymmetry of optimality theoretic syntax and semantics. *Technical report*. University of Amsterdam.

Zeevat, H.: 1999b. Old material. Handout for talk at the ESSLLI-99 workshop on Focus and Presupposition in Multi-Speaker Discourse.

Zeevat, H.: 1999c. Semantics in optimality theory. *in* H. de Hoop and H. de Swart (eds), *OTS Papers on Optimality Theoretic Semantics*. OTS, Utrecht University. Utrecht.

Zimmermann, E.: 1999. Remarks on the epistemic role of discourse referents. *in* L. S. Moss, J. Ginzburg and M. de Rijke (eds), *Logic, Language and Computation, Vol II*. Kluwer. Stanford, CA.

Zuidema, M.: 1999. *The culprit might have done it*. Master's thesis. Utrecht University.

Samenvatting

Iemand heeft Spiderman vermoord, na uitgebreid onderzoek ontdek je dat John Smith de dader is en nu wil je hem arresteren. Hij is op een gemaskerd bal. Je gaat erheen maar je weet niet hoe hij eruitziet. Is de zin 'Je weet wie Spiderman vermoord heeft' nu waar of niet waar? Aan de ene kant is de zin waar, je weet dat John Smith het gedaan heeft. Aan de andere kant is de zin ook onwaar. Omdat je niet weet hoe hij eruitziet kan je hem niet aanwijzen: deze persoon kan de dader zijn, maar ook die persoon daar. De evaluatie van deze zin blijkt afhankelijk te zijn van de manier waarop je individuen identificeert. Dat kan op meerdere manieren, zoals door middel van naamgeving (John Smith, Bill White, enzovoort), of door aan te wijzen (deze persoon hier, die persoon daar, enzovoort). Als je uitgaat van identificatie door naamgeving is de zin waar, als je uitgaat van identificatie door aanwijzing is de zin onwaar.

Dit voorbeeld illustreert de centrale stelling die ik verdedig in dit proefschrift. In verschillende conversationele omstandigheden worden verschillende identificatie-methoden gebruikt en de evaluatie van tekstfragmenten in natuurlijke taal variëert met de gebruikte methodes. Klassieke semantische theoriën abstraheren van de specifieke manieren waarop individuen geïdentificeerd worden en kunnen daarom niet overweg met deze afhankelijkheid. Met de analyse die ik voorstel wordt wel rekening gehouden met de verschillende identificatie-methoden en kunnen we een verantwoording geven van hun invloed op de interpretatie.

Vragen, geloofstoekenningen en epistemisch modale uitdrukkingen zijn voorbeelden van constructies waarvan de interpretatie afhangt van de manier waarop objecten geïdentificeerd worden. In dit proefschrift zal ik deze drie constructies bestuderen uitgaande van de zogeheten 'partitietheorie' van vragen, van de modale predikaten logica voor geloofstoekenningen, en van een intensionele dynamische semantiek voor epistemische modale uitdrukkingen. De drie theorieën maken gebruik van de notie van een *mogelijke wereld.* Mogelijke werelden zijn theoretische evaluatiepunten waar de lexicale uitdrukkingen van een taal een (mogelijk variërende) interpretatie hebben. Een mogelijke wereld geeft een voor mogelijk gehouden constellatie van feiten weer, in termen van de (atomaire) zinnen die als waar gelden in die wereld. Mogelijke werelden kunnen aldus dienen om de kennis of het geloof van iemand te karakteriseren, in termen van die deelverzamelingen die verenigbaar zijn met wat die persoon weet of gelooft. In de interpretatie van vragen, geloofstoekenningen, en epistemische modalen kunnen bepaalde operatoren een overstap inhouden van de ene evaluatiewereld naar een andere. Operatoren met deze eigenschap worden *intensioneel* genoemd.

De context gevoelige constructies die ik zal behandelen worden in de klassieke literatuur gerepresenteerd met behulp van logische formules waarin een variabele vrij in het syntactisch bereik van een intensionele operator voorkomt. In normale logische systemen bestaat het interpretatie domein van variabelen uit 'kale' individuen, en daarom wordt geen rekening gehouden met de wijze deze individuen geïdentificeerd worden, en met de contextafhankelijkheid van deze identificatiemethodes.

Mijn analyse houdt de klassieke representatie voor dit type zinnen in stand, maar kent een andere betekenis toe aan variabelen in een intensionele context. Voor een deel houdt mijn voorstel in variabelen te interpreteren als functies van werelden naar objecten in plaats van de objecten zelf. Deze functies worden traditioneel *individuele concepten* genoemd. Individuele concepten vormen een manier om objecten te identificeren in verschillende mogelijke werelden. Het andere deel van mijn voorstel houdt in kwantoren te interpreteren over domeinen van concepten die (a) contextueel gedetermineerd zijn en (b) voldoen aan de volgende eis: voor iedere wereld is ieder individu geïdentificeerd door één en slechts één concept in de contextueel relevante verzameling. Ik noem verzamelingen van concepten die aan deze eis voldoen conceptuele bedekkingen *(conceptual covers)*. Een conceptuele bedekking vormt een methode om de individuen in een domein te identificeren, en verschillende conceptuele bedekkingen komen dus overeen met verschillende manieren om dat domein te zien. Door de kwantoren in de drie bovengenoemde theorieën te relateren aan verschillende conceptuele bedekkingen, kan de interpretatie van vragen, geloofstoekenningen, en epistemische modale uitdrukkingen afhankelijk gemaakt worden van de pragmatisch actieve conceptualisaties van het gespreksdomein *(universe of discourse)*. Met deze relativering kunnen we een aantal traditionele en nieuwe problemen oplossen die in deze contexten rijzen; tegelijkertijd kunnen we de specifieke problemen voorkomen die normaal gesproken ontstaan wanneer we kwantificeren over concepten in plaats van over objecten.

Organisatie van het proefschrift

De eerste drie hoofdstukken van dit proefschrift kunnen onafhankelijk van elkaar gelezen worden. De hoofdstukken 1, 2, en 3 zijn ontstaan als onafhankelijke artikelen, geschreven gedurende verschillende perioden van mijn promotiestudie. Hoofdstuk 4 bouwt voort op de eerdere drie, en staat dus niet op zichzelf.

Hoofdstuk 1 behandelt de interpretatie van vragen en vraagwoord-constructies. Het is ontstaan uit materiaal dat ik gepresenteerd heb in Leipzig (Sinn und Bedeutung 1998) en Stanford (LLC 1999). In dit hoofdstuk presenteer ik een verfijning van de logica van vragen van Groenendijk & Stokhof die het mogelijk maakt zulke constructies afhankelijk te maken van specifieke conceptualisaties van het gespreksdomein. Ik laat zien dat er op deze manier een aantal fundamentele problemen voorkomen kunnen worden. Daarna breid ik mijn analyse uit tot twee andere taalkundige theorieën van vragen, de verzamelingen proposities theorie, en de gestructureerde betekenis theorie. Een deel van hoofdstuk 1 zal verschijnen als Aloni (to appear).

In hoofdstuk 2 analyseer ik de interpretatie van propositionele attitudes, in het bijzonder van geloofstoeschrijvingen. Het hoofdstuk is ontstaan uit Aloni (1998). In het eerste deel bespreek ik klassieke puzzels betreffende de wisselwerking tussen propositionele attitudes, kwantoren en het concept van identiteit. Ik vergelijk verschillende reacties op deze puzzels binnen het raamwerk van de modale predikaten logica en pleit voor een analyse waarin *de re* geloofstoekenningen mede afhankelijk worden gemaakt van de methodes om objecten te identificeren die geschikt zijn in de gebruiksomstandigheden. In het tweede deel van het hoofdstuk geef ik deze analyse een precieze formalisatie, en presenteer een modale predikaten logica met conceptuele bedekkingen vanuit een model- en een bewijs-theoretisch perspectief. Het resulterende systeem wordt vergeleken met normale modale predikaten logica en ik behandel een aantal toepassingen.

Hoofdstuk 3 behandelt de combinatie van dynamische kwantoren met 'holistische noties' zoals epistemische modaliteit, presuppositie en dynamische ondersteuning. Ik vergelijk de verschillende wijzen waarop dynamische kwantificatie gedefiniëerd kan worden, beargumenteer dat deze allemaal empirische en theoretische moeilijkheden opleveren in combinatie met de genoemde holistische noties. Daarna toon ik aan dat kwantificatie onder conceptuele bedekkingen deze moeilijkheden voorkomt. Het hoofdstuk komt voort uit Aloni (1997a) en Aloni (1997b). Voor het meerendeel is het reeds gepubliceerd als Aloni (2000).

Hoofdstuk 4 onderzoekt formele en pragmatische aspecten van conceptuele bedekkingen. In het eerste deel bestudeer ik eerst een aantal formele eigenschappen van conceptuele bedekkingen, en vergelijk vervolgens de notie van identificatie onder een conceptuele bedekkingen met andere vormen van identificatie van individuen in verschillende werelden. In het tweede deel bediscussiëer ik de pragmatische selectie van conceptuele bedekkingen. De selectie van bedekkingen is afhankelijk van de context en blijkt bepaald te worden door interpretatie- en generatie-regels die 'zacht' zijn, in de terminologie van optimaliteitstheorie. Ik schets de contouren van een twee-dimensionale optimaliteitstheoretische interpretatie die is geformuleerd met behulp van speltheoretische concepten. Speltheorie blijkt hier een veelbelovend raamwerk te zijn voor het bestuderen van de afwegingen die een spreker en zijn gehoor maken in hun zoektocht naar een optimale interpretatie van context gevoelige uitdrukkingen in de natuurlijke taal.

Titles in the ILLC Dissertation Series:

ILLC DS-1996-01: **Lex Hendriks**
Computations in Propositional Logic

ILLC DS-1996-02: **Angelo Montanari**
Metric and Layered Temporal Logic for Time Granularity

ILLC DS-1996-03: **Martin H. van den Berg**
Some Aspects of the Internal Structure of Discourse: the Dynamics of Nominal Anaphora

ILLC DS-1996-04: **Jeroen Bruggeman**
Formalizing Organizational Ecology

ILLC DS-1997-01: **Ronald Cramer**
Modular Design of Secure yet Practical Cryptographic Protocols

ILLC DS-1997-02: **Nataša Rakić**
Common Sense Time and Special Relativity

ILLC DS-1997-03: **Arthur Nieuwendijk**
On Logic. Inquiries into the Justification of Deduction

ILLC DS-1997-04: **Atocha Aliseda-LLera**
Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence

ILLC DS-1997-05: **Harry Stein**
The Fiber and the Fabric: An Inquiry into Wittgenstein's Views on Rule-Following and Linguistic Normativity

ILLC DS-1997-06: **Leonie Bosveld - de Smet**
On Mass and Plural Quantification. The Case of French 'des'/'du'-NP's.

ILLC DS-1998-01: **Sebastiaan A. Terwijn**
Computability and Measure

ILLC DS-1998-02: **Sjoerd D. Zwart**
Approach to the Truth: Verisimilitude and Truthlikeness

ILLC DS-1998-03: **Peter Grunwald**
The Minimum Description Length Principle and Reasoning under Uncertainty

ILLC DS-1998-04: **Giovanna d'Agostino**
Modal Logic and Non-Well-Founded Set Theory: Translation, Bisimulation, Interpolation

ILLC DS-1998-05: **Mehdi Dastani**
Languages of Perception

ILLC DS-1999-01: **Jelle Gerbrandy**
Bisimulations on Planet Kripke

ILLC DS-1999-02: **Khalil Sima'an**
Learning efficient disambiguation

ILLC DS-1999-03: **Jaap Maat**
Philosophical Languages in the Seventeenth Century: Dalgarno, Wilkins, Leibniz

ILLC DS-1999-04: **Barbara Terhal**
Quantum Algorithms and Quantum Entanglement

ILLC DS-2000-01: **Renata Wasserman**
Resource Bounded Belief Revision

ILLC DS-2000-02: **Jaap Kamps**
A Logical Approach to Computational Theory Building (with applications to sociology)

ILLC DS-2000-03: **Marco Vervoort**
Games, Walks and Grammars: Problems I've Worked On

ILLC DS-2000-04: **Paul van Ulsen**
E.W. Beth als logicus

ILLC DS-2000-05: **Carlos Areces**
Logic Engineering. The Case of Description and Hybrid Logics

ILLC DS-2000-06: **Hans van Ditmarsch**
Knowledge Games

ILLC DS-2000-07: **Egbert L.J. Fortuin**
Polysemy or monosemy: Interpretation of the imperative and the dative-infinitive construction in Russian