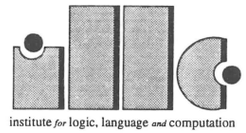
Proceedings of the Tenth Amsterdam Colloquium

December 18 — 21, 1995

Part III

Paul Dekker and Martin Stokhof
(eds.)

ILLC/Department of Philosophy
University of Amsterdam



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

Proceedings of the

Tenth Amsterdam Colloquium

Printing by CopyPrint 2000, Enschede

Cover design by Instand, Malden

ISBN: 90-74795-52-8

Contents

Presuppositional Discourse Representation Theory

Emiel Krahmer

1 Introduction

In this paper we will study presuppositions from the perspective of *Discourse Representation Theory* (DRT). To do this we will extend the system of Double Negation DRT, discussed in Krahmer and Muskens (1994, 1995), with presuppositional representations. The resulting system will be referred to as *Presuppositional DRT*. As far as the *interpretation* is concerned, Presuppositional DRT will turn out to be closely related to Beaver's Kinematic Predicate Logic and van Eijck's Error-state Semantics for DPL (both incorporating a form of partiality in a dynamic semantic framework),[1] although we shall propose a different treatment of negation and disjunction. On the other hand, the *representations* of Presuppositional DRT will show great resemblance to those employed by van der Sandt in his *presuppositions-as-anaphora* theory.[2]

Thus, Presuppositional DRT may be associated with *two* different approaches to presuppositions: a semantic one in the tradition of Karttunen, Heim, Beaver, van Eijck and others, and an algorithmic one in the style of van der Sandt's presuppositions-as-anaphora approach. The resulting picture obviously enhances comparison between the two approaches, but in this paper we will also discuss an additional advantage: taking Presuppositional DRT as an alternative foundation improves the presuppositions-as-anaphora theory. We will not propose any significant *changes* to van der Sandt's theory, rather we will show that merely defining van der Sandt's presupposition resolution algorithm on top of Presuppositional DRT instead of standard DRT has a number of advantages.[3]

The rest of this paper contains the following sections. In section 2 we discuss van der Sandt's theory in more detail. In sections 3, 4 and 5 we present Presuppositional DRT (the formal system, the applications and a method to calculate the semantic presupposition of arbitrary DRSs). In section 6 we re-discuss van der Sandt's theory, but now in terms of Presuppositional DRT. In the two final sections some of the questions raised by the present perspective are discussed.

2 Presuppositions as anaphora

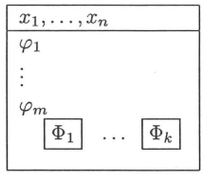
In the late eighties, Van der Sandt observed that there is a striking correspondence between the behavior of anaphora in discourse and the projection of presuppositions in complex sentences (Sandt (1989)). This observation is the crux of the *presuppositions-as-anaphora*-theory, in which it is proposed to *resolve* presuppositions just like we resolve anaphoric pronouns in DRT. Van der Sandt formulates the presuppositions-as-anaphora approach in terms of a modified version of DRT, which differs from standard DRT only in the way DRSs are constructed. In standard DRT, the construction algorithm turns (the syntactic tree of) a sentence into a DRS, and pronouns are resolved *en route*. In the presuppositions-as-anaphora theory a curtailed version of the construction algorithm (which does not resolve anaphora) produces a so-called *S-DRS* for a sentence, syntactically defined as follows:

[1] See Beaver (1992, 1993, 1994, 1995) and Eijck (1993a, 1993b, 1994) respectively.
[2] See Sandt and Geurts (1991) and Sandt (1992), and the extension in e.g. Geurts (1994).
[3] In spirit, Presuppositional DRT can also be compared with the system of Zeevat (1992). Zeevat proposes a reformulation of van der Sandt's theory in terms of Update Semantics and compares this reformulation with other approaches to presupposition, in particular the one found in Heim (1983). But where Zeevat re-defines van der Sandt's theory in the formalism of Update Semantics, we will leave van der Sandt's representational theory essentially untouched. An alternative approach can be found in Fernando (1994), where *syntactic translations* are used to produce the presuppositions of dynamic formulae.

If $x_1, \ldots, x_n$ are variables, $\varphi_1, \ldots, \varphi_m$ are conditions, and $\Phi_1, \ldots, \Phi_k$ are S-DRSs, then $[x_1, \ldots, x_n \mid \varphi_1, \ldots, \varphi_m \mid \Phi_1, \ldots, \Phi_k]$ is an S-DRS, $n, m, k \geq 0$ (Sandt (1992:354)).

In a pictorial format an S-DRS looks as follows:

(DRS 1)



So, an S-DRS can be seen as a standard DRS with embedded (S-)DRSs Φ_i added, each representing an elementary presupposition. Notice that since the presuppositional Φ_i are S-DRSs as well, they may themselves contain embedded presuppositions. In this section we will often use the term 'DRS' to refer to an S-DRS, trusting that this will not lead to confusion.

2.1 Resolving presuppositions

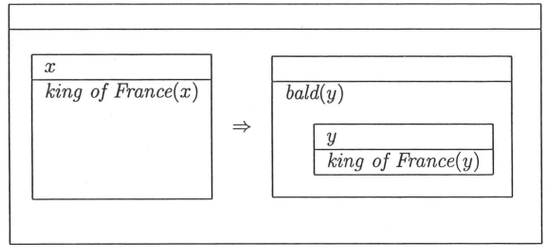
After an S-DRS has been constructed, the presuppositional representations (if any) are resolved in the context of the representation of the discourse so far. In van der Sandt's theory, resolving a presupposition amounts to either *binding* or *accommodating* it. For this purpose, a *presupposition resolution algorithm* is developed, consisting of three consecutive stages, beginning with:[4]

> ONE: *Try to* bind *the presuppositional DRS as low as possible.*

Since it is assumed that presupposition is just a species of anaphora, we start with looking for a *suitable, accessible* antecedent. As usual, accessibility is determined by the structure of the relevant DRS. But now we can also say something about the *suitability* of the antecedent: it has to satisfy the conditions of the presuppositional DRS.[5] Consider (1) and van der Sandt's representation (DRS 2).

(1) If France has a king, then the king of France is bald.

(DRS 2)



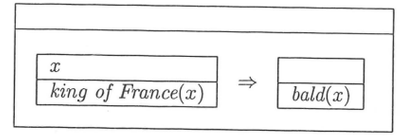
Here the consequent consists of two DRSs: one embedded DRS representing the presupposition triggered in the consequent of the implication in (1) ('there is a king of France') and one representing the assertion ('he is bald'). To *resolve* the presupposition, we look for an accessible, suitable antecedent, and obviously we

[4] For the formal details, see any of the references given in footnote 2.

[5] In fact, van der Sandt also allows for *partial matches* between presupposition and antecedent, but for now we will ignore this possibility.

find one in the antecedent-DRS, namely x. So, we can *bind* the presuppositional DRS: the representation of the elementary presupposition is taken away from the consequent-DRS and 'moved into' the antecedent-DRS. To round things off, all free occurrences of y are replaced by its newly found antecedent, x:

(DRS 3)



Notice that (DRS 3) is also the standard DRT representation for example (2). In fact, the parallel between examples like (1) and (2) is one of the suggestive facts discussed by van der Sandt to motivate his presuppositions-as-anaphora approach.

(2) If France has a king, then he is bald.

In this case, we easily succeeded in finding an antecedent, and hence in binding the presupposition: there was only one possibility. If there is more than one possible antecedent for a presupposition, van der Sandt's algorithm prefers binding to the *lowest* antecedent, i.e. the closest accessible discourse referent.

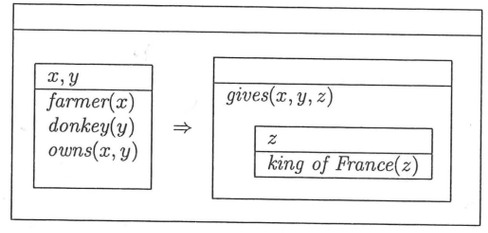
It is also possible that we have checked every accessible DRS without finding a possible antecedent for the presupposition, not even on the highest possible landing-site: the main DRS. Then we go the next phase of the resolution algorithm:

> TWO: *If binding a presuppositional DRS fails, try to* accommodate *it as high as possible.*

Accommodation of a presuppositional DRS Φ in some DRS Ψ simply amounts to adding Φ to Ψ, i.e. we replace Ψ with the merge of the two DRSs Φ and Ψ. When trying to accommodate a presuppositional DRS we first try to accommodate in the main DRS ('global accommodation'), before we attempt accommodation in lower (sub-)DRSs ('local accommodation'). Consider example (3), with its DRT representation (DRS 4).

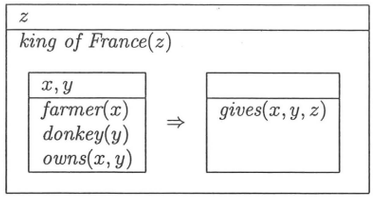
(3) If a farmer owns a donkey, he gives it to the king of France.

(DRS 4)



To be precise, this DRS is already partially resolved; the representations of the pronouns *he* and *it* have been bound to the antecedents *a farmer* and *a donkey* respectively. One presuppositional DRS remains, expressing the elementary presupposition that there is a king of France. Obviously, binding it is out of the question, since none of the accessible discourse referents satisfies the condition that it is a king of France. So, we attempt to accommodate the presupposition. That is: we remove the presuppositional DRS $[z \mid \textit{king of France}(z)]$ from the consequent, and merge it with what is left of (DRS 4) after the removal of the presuppositional DRS, which gives the following.

(DRS 5)



Accommodation is a very strong mechanism, and therefore van der Sandt requires the result of accommodation to satisfy certain constraints. One obvious constraint is that variables may not end up being free after accommodation. Moreover, the result of accommodation must satisfy certain general, independently motivated constraints: it has to be both *informative* and *consistent*. Roughly, accommodating Ψ in Φ with Φ' as output is (i) informative if the set of models which support Φ' is a proper subset of the set of models supporting Φ and (ii) consistent if there is at least one model satisfying Φ'. If accommodating a presuppositional DRS at a certain level violates one or more of the constraints, we try to accommodate it on a lower level. If no accommodation is possible, we come to the last phase:

THREE: *If both binding and accommodation fail: Give up*

This would happen with an example like (4):

(4) There is no king of France. ?? The king of France is bald.

If the reader constructs a mental DRS for this example, she may easily verify that binding the presupposition that there is a king of France (triggered by the definite description in the second sentence) will not succeed since there are no accessible antecedents, and similarly any accommodation will violate the consistency condition.

Van der Sandt's presuppositions-as-anaphora theory has several highly attractive features. It uses a single mechanism for presuppositions and 'ordinary' anaphora, and it has an integrated analysis which treats binding and accommodation as two sides of the same coin. Moreover, on the empirical side it does remarkably well. In general, it is fair to say that there is no other theory (static or dynamic) which deals in a successful way with quite the same range of data as the presuppositions-as-anaphora approach. Nevertheless it raises a few questions as well.

2.2 What is a presuppositional DRS?

To begin with, van der Sandt modifies the syntax of DRT so as to include presuppositional DRSs in it. But, there is no interpretation for these new S-DRSs. One can argue that representations such as (DRS 2) are in fact *proto DRSs*; i.e. they are still under construction. To some extent this position can be defended: the representation *is* under construction since some anaphoric elements are still to be resolved. And, when they are all resolved, the net-result is a standard DRS which has a perfectly normal interpretation. On the other hand: we would like to have an interpretation for *any* expression of the DRS language. By comparison, the proto-DRSs which occur in Kamp and Reyle (1993) are *not* elements of the DRT language; they are just partly dismantled syntactic trees with a box around them. So we have two options: either representations like (DRS 2) are indeed proto-DRSs, and hence they are not elements of the proper DRT language (and thus do not require interpretation). Or, we do not treat them like proto-DRSs but like ordinary DRSs, and then they do require an interpretation. Here we will argue for the second option. We shall see that Presuppositional DRT can be used to give an interpretation to S-DRSs. Moreover, it will be shown that having such an interpretation actually has several positive effects.

2.3 Procedural vs. declarative

For example, van der Sandt has observed that his (procedural) way of resolving presuppositions we just sketched '(...) *obviously leads to problems when we check for logical properties such as consistency or entailment. For interpretation and determination of these logical properties can only ensue after full resolution of all anaphoric expressions.*' (Sandt (1992:362)). After all, S-DRSs do not have an interpretation, ordinary DRSs without embedded presuppositions do. Hence in more complex cases, the procedural version of the algorithm will involve '*a substantial amount of backtracking*' (Sandt (1992:*ibid.*)). Therefore, van der Sandt opts for a more declarative approach: for an unresolved S-DRS Φ a set of all *logically possible* resolutions for Φ is constructed. That is, every possible resolution is carried out and interpreted. By the underlying principles we discussed above (binding as low as possible, accommodation as high as possible, binding preferred over accommodation), this set of logical possibilities is partially ordered, and the resolution which occupies the first place in the ordering —the preferred reading— will also be the output of the procedural version of the resolution algorithm. For simple sentences the difference between the procedural and the declarative approach is not very important, but when considering longer stretches of text the set of all logically possible resolutions may turn out to be rather big, and computing all of them may take some time. If there is an interpretation for presuppositional DRSs we can check the logical properties on the spot, and the procedural algorithm will not run into problems; we can stop when we have found *one* satisfactory resolution.[6] Of course this does not mean that we *have* to stop: if we wanted to, we could continue searching for less preferred resolutions as long as we like.

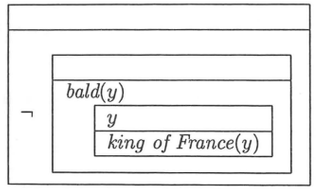
2.4 Accommodating failing presuppositions

Consider example (5).

(5) It is not the case that the king of France is bald.

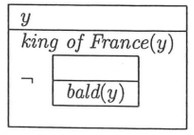
On van der Sandt's approach this sentence will be represented by (DRS 6).

(DRS 6)



Once this representation is constructed we try to resolve the presuppositional DRS. Binding it is not possible, so we accommodate the presupposition in the main DRS, which obviously does not violate any constraints. The result is depicted in (DRS 7).

(DRS 7)

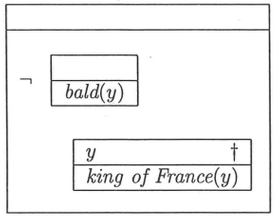


Now suppose we interpret (DRS 7) in a model in which France has no king. In such a model (DRS 7) is false; here van der Sandt essentially follows the Russellian scoping

[6] Although having an interpretation for presuppositional DRSs leads to a substantial reduction of the search space, some amount of backtracking is inevitable. See footnote 18 below.

analysis. Kracht (1994) objects to this way of accommodating presuppositions (or 're-allocating' them, as Kracht calls it) since he observes that an accommodated presupposition should still be a presupposition. However, in van der Sandt's theory a presupposition loses its *presuppositionhood* when it is resolved. Notice that there is no other option since presuppositional DRSs do not have an interpretation. Interestingly, van der Sandt remarks that we can restore the Frege/Strawson intuition that a failing presupposition leads to undefinedness '(...) *by explicitly marking accommodated material as such and making the embedding function dependent on the status of the relevant markers*' (Sandt (1992:375)). In other words: accommodated, presuppositional material has to be distinguished from simply asserted material. When we define van der Sandt's theory on top of Presuppositional DRT instead of standard DRT, we can easily mark something as presuppositional, namely by representing it as a presuppositional DRS. This would call for a small change in the resolution algorithm: we accommodate the presuppositional DRS *as a presuppositional DRS*. Thus, resolving the presuppositional DRS in (DRS 6) does not result in (DRS 7), but in (DRS 8). To indicate that the presuppositional DRS has been 'resolved' it is marked with a †.[7]

(DRS 8)



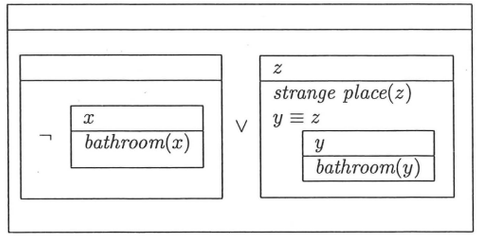
As we shall see below the interpretation of a presuppositional DRS is defined is such a way that in a model where there is no king of France, (DRS 8) indeed will be undefined, as Frege and Strawson would have it.

2.5 *Disjunctions*

Consider the following (lexical variant of an) example discussed in Sandt (1992:368) and its representation (DRS 9).

(6) Either there is no bathroom in this house, or the bathroom is in a strange place.

(DRS 9)



The presuppositions-as-anaphora theory, based on standard DRT, predicts that binding the presuppositional DRS $[y \mid bathroom(y)]$ is not possible, since in standard DRT a (negated) left disjunct is not accessible for a right disjunct. Global accommodation of the presupposition triggered by *the bathroom* will not work either, since it would represent the unacceptable discourse *There is a bathroom in this house. Either there is no bathroom in this house, or it is in a strange place*; it violates the consistency constraint for subDRSs. As a consequence, we have to go for the least preferred way of resolving the presuppositional DRS: it is locally accommodated in the disjunct it originated in. The resulting reading may be paraphrased as *either there is no bathroom in this house, or there is a bathroom in this house and it is in a strange place*. It is questionable whether this really captures the meaning of sentence (6). For instance, what if there are *two* bathrooms in this house, one is and is not in a strange place? Be that as it may, Beaver (1995) shows that things go wrong when considering the following, comparable example:

(7) Either Mary's autobiography hasn't appeared yet, or else John must be very proud that Mary has had a book published.

Let us assume that *being proud that S* presupposes S, that is: the second disjunct presupposes that Mary has had a book published. Van der Sandt's theory predicts that the most preferred reading of (7) is the one involving global accommodation. Binding is again not possible, but the global accommodation of *Mary has had a book published* is not ruled out. Nevertheless, we would want to analyze this example in terms of binding; on its most natural reading, (7) does *not* presuppose that Mary has had a book published. Below we will see that Presuppositional DRT helps.

2.6 *Presupposition-quantification interaction*

Van der Sandt's theory has been criticized for its treatment of presuppositions in the scope of quantifiers (see for instance Beaver (1995)). A typical example is (8), which is a variant of an example discussed in Heim (1983):

(8) Every German loves his car.

Van der Sandt's resolution algorithm predicts that the presupposition triggered by *his car* is accommodated in the scope of *every*. It cannot be accommodated in the main DRS because the referent of *his* would end up being free, which is forbidden by the so-called *trapping* constraint. Since the presupposition loses its presuppositionhood it is predicted that (8) is equivalent with (9).

(9) Every German who has a car loves it.

The question what the exact meaning is of sentences like (8) is hard to decide (see e.g. Beaver (1994) for discussion), but the alleged equivalence between (8) and (9) is not generally supported. As we shall see, Presuppositional DRT may shed a new light on these examples. Now, let us first discuss the system.

3 Presuppositional DRT

Presuppositional DRT differs in three significant respects from ordinary DRT. First of all, it embodies a different treatment of negation and disjunction due to Krahmer and Muskens (1994, 1995). The new negation validates the law of double negation, while it leaves the usual DRT analysis of single negations untouched. The basic problem with negation in standard DRT is that it is not a 'flip-flop' operation like negation in ordinary logic. To see this we only have to look at the DRT syntax: if Φ is a DRS, $\neg\Phi$ is a condition, and there is no comparable operator which takes us from conditions to DRSs again. To remedy this, the new negation $\sim\Phi$ of a DRS Φ is itself a DRS. Disjunction is redefined in terms of this new negation, which will be seen to enhance the treatment of the notorious bathroom-sentences.

[7] Below we shall see that the free occurrence of y in the condition $bald(y)$ is bound by the introduction of y in the presuppositional DRS.

A second difference with ordinary DRT is that Presuppositional DRT contains representations for presuppositional DRSs, which are designated as $\partial\Phi$ (this unary operator is due to Beaver (1992)). The intuition behind $\partial\Phi$ is that it says of Φ that it is an elementary presupposition. $\partial\Phi$ will be interpreted in such a way that it is supported (True) whenever Φ is supported (True), and neither supported nor rejected otherwise. In particular: $\partial\Phi$ cannot be rejected (False). This exhausts the syntactic differences with standard DRT so that the syntax of Presuppositional DRT is defined as follows. Let CON be the (non-empty) set of constants, while VAR is the (similarly non-void) set of discourse referents/variables. The union of these two gives the set of terms.

DEFINITION 1 (Presuppositional DRT syntax)

1. If R is an n-ary predicate and $t_1, \ldots, t_n$ are terms, then $R(t_1, \ldots, t_n)$ is a condition.
2. If t_1 and t_2 are terms, then $t_1 \equiv t_2$ is a condition.
3. If Φ and Ψ are DRSs, then $(\Phi \Rightarrow \Psi)$ is a condition.
4. If $x_1, \ldots, x_n$ are variables and $\varphi_1, \ldots, \varphi_m$ are conditions (with $n, m \geq 0$), then $[x_1, \ldots, x_n \mid \varphi_1, \ldots, \varphi_m]$ is a DRS.
5. If Φ and Ψ are DRSs, then $(\Phi\,;\Psi)$, $\sim\Phi$ and $\partial\Phi$ are DRSs.

The third difference with the usual conception of DRT is that Presuppositional DRT has a partial interpretation: each condition and DRS will be associated with *two* interpretations: a positive interpretation (designated as $[\![.]\!]^+$) and a negative one ($[\![.]\!]^-$). Let F be the set of finite assignments, with Λ as the empty assignment. Here and elsewhere we will write $\vec{x}$ for $x_1, \ldots, x_n$. $g\{\vec{x}\}h$ abbreviates: assignment h extends assignment g with $\vec{x}$.[8] Let $M = \langle D, I\rangle$ be an ordinary first-order model. For terms the interpretation is defined as follows: $[\![t]\!]^+_{M,g} = [\![t]\!]^-_{M,g} = g(t)$ if $t \in$ VAR and $t \in$ DOM(g), and $[\![t]\!]^+_{M,g} = [\![t]\!]^-_{M,g} = I(t)$ if $t \in$ CON. Since for any term t, $[\![t]\!]^+$ equals $[\![t]\!]^-$, we will drop the superscript in the interpretation of terms. If $g(x)$ is not defined, $[\![x]\!]_g$ is not defined either. For conditions φ, we define $[\![\varphi]\!]^+_M \subseteq F$ and $[\![\varphi]\!]^-_M \subseteq F$, while for DRSs Φ we define $[\![\Phi]\!]^+_M \subseteq F^2$ and $[\![\Phi]\!]^-_M \subseteq F^2$ as follows (in the first clause $1 \leq i \leq n$, we drop the M-index where this is possible without confusing the reader):[9]

DEFINITION 2 (Presuppositional DRT semantics)

1. $[\![R(t_1, \ldots, t_n)]\!]^+ = \{g \mid [\![t_i]\!]_g$ *defined* $\& \langle [\![t_1]\!]_g, \ldots, [\![t_n]\!]_g\rangle \in I(R)\}$
   $[\![R(t_1, \ldots, t_n)]\!]^- = \{g \mid [\![t_i]\!]_g$ *defined* $\& \langle [\![t_1]\!]_g, \ldots, [\![t_n]\!]_g\rangle \notin I(R)\}$
2. $[\![t_1 \equiv t_2]\!]^+ = \{g \mid [\![t_1]\!]_g, [\![t_2]\!]_g$ *defined* $\& [\![t_1]\!]_g = [\![t_2]\!]_g\}$
   $[\![t_1 \equiv t_2]\!]^- = \{g \mid [\![t_1]\!]_g, [\![t_2]\!]_g$ *defined* $\& [\![t_1]\!]_g \neq [\![t_2]\!]_g\}$
3. $[\![\Phi \Rightarrow \Psi]\!]^+ = \{g \mid \forall h(\langle g, h\rangle \in [\![\Phi]\!]^+ \Rightarrow \exists k\langle h, k\rangle \in [\![\Psi]\!]^+) \;\&\; \mathrm{DEF}_g(\Phi)\}$
   $[\![\Phi \Rightarrow \Psi]\!]^- = \{g \mid \exists h(\langle g, h\rangle \in [\![\Phi]\!]^+ \;\&\; \exists k\langle h, k\rangle \in [\![\Psi]\!]^-)\}$

[8] Formally: DOM(h) = DOM$(g) \cup \{\vec{x}\}$ $\&$ $\forall y \in$ DOM$(g) : g(y) = h(y)$.

[9] The interpretation of Presuppositional DRT derives from the second middle Kleene based dynamic interpretation discussed in Krahmer (1994), definition 20. As said above, the interpretation is closely related to the partial dynamic systems developed for the treatment of presuppositions by Beaver and van Eijck (see the references mentioned in footnote 1). This is clearest for the system given in Eijck (1994:779-780). The main difference is that Presuppositional DRT is based on Double Negation DRT, and that as a consequence negation and disjunction are treated in a different way. Other combined partial dynamic interpretations, devised for different purposes, can be found in e.g. Dekker (1993), Piwek (1993) and van den Berg (1991,1993).

4. $[\![[\vec{x} \mid \varphi_1, \ldots, \varphi_m]]\!]^+ = \{\langle g, h\rangle \mid g\{\vec{x}\}h \;\&\; h \in ([\![\varphi_1]\!]^+ \cap \ldots \cap [\![\varphi_m]\!]^+)\}$
   $[\![[\vec{x} \mid \varphi_1, \ldots, \varphi_m]]\!]^- = \{\langle g, g\rangle \mid \forall h(g\{\vec{x}\}h \Rightarrow h \in ([\![\varphi_1]\!]^- \cup \ldots \cup [\![\varphi_m]\!]^-))\}$
5. $[\![\Phi\,;\Psi]\!]^+ = \{\langle g, h\rangle \mid \exists k(\langle g, k\rangle \in [\![\Phi]\!]^+ \;\&\; \langle k, h\rangle \in [\![\Psi]\!]^+)\}$
   $[\![\Phi\,;\Psi]\!]^- = \{\langle g, g\rangle \mid \forall k(\langle g, k\rangle \in [\![\Phi]\!]^+ \Rightarrow \exists h\langle k, h\rangle \in [\![\Psi]\!]^-) \;\&\; \mathrm{DEF}_g(\Phi)\}$
6. $[\![\sim\Phi]\!]^+ = [\![\Phi]\!]^-$
   $[\![\sim\Phi]\!]^- = [\![\Phi]\!]^+$
7. $[\![\partial\Phi]\!]^+ = [\![\Phi]\!]^+$
   $[\![\partial\Phi]\!]^- = \emptyset$

Where $\mathrm{DEF}_{M,g}(\Phi)$ ('DRS Φ is dynamically defined in M with respect to g') is defined as follows (again suppressing M).

DEFINITION 3 (DEF)

$\mathrm{DEF}_g(\Phi) = \exists h\langle g, h\rangle \in [\![\Phi]\!]^+$ or $\exists h\langle g, h\rangle \in [\![\Phi]\!]^-$ (for Φ is $\partial\Psi$ or $[\vec{x} \mid \varphi_1, \ldots, \varphi_m]$)
$\mathrm{DEF}_g(\Phi\,;\Psi) = \mathrm{DEF}_g(\Phi) \;\&\; \forall h(\langle g, h\rangle \in [\![\Phi]\!]^+ \Rightarrow \mathrm{DEF}_h(\Psi))$
$\mathrm{DEF}_g(\sim\Phi) = \mathrm{DEF}_g(\Phi)$

We will say that a DRS Φ is supported in a model M with respect to an assignment g (notation: $M, g \models \Phi$) iff $\exists h\langle g, h\rangle \in [\![\Phi]\!]^+_M$, and that Φ is rejected in M with respect to g (notation: $M, g \dashv \Phi$) iff $\exists h\langle g, h\rangle \in [\![\Phi]\!]^-_M$. The truth-combinations can now be defined as follows. Let Φ be a proper DRS, then:[10]

DEFINITION 4 (Truth-combinations)

Φ is *True* in M iff $M, \Lambda \models \Phi$
Φ is *False* in M iff $M, \Lambda \dashv \Phi$
Φ is *Neither* in M iff $M, \Lambda \not\models \Phi$ and $M, \Lambda \not\dashv \Phi$

It is easily shown that no DRS is both True and False, and that every DRS which does not contain presuppositions is either True or False. Two DRSs Φ and Ψ are said to be equivalent iff for all models M: $[\![\Phi]\!]^+_M = [\![\Psi]\!]^+_M$ and $[\![\Phi]\!]^-_M = [\![\Psi]\!]^-_M$. We introduce the following abbreviations.[11]

DEFINITION 5 (Abbreviations)
$\Phi_{\langle\Psi\rangle}$ abbreviates $\partial(\Psi)\,;\Phi$
$\Phi \vee \Psi$ abbreviates $\sim\Phi \Rightarrow \Psi$

The first abbreviation defines (a dynamic version of) Blamey (1986)'s transplication in terms of Beaver's unary presupposition operator, which is our linear alternative for the presuppositional DRSs. It results in the following interpretation for $\Phi_{\langle\Psi\rangle}$:

$[\![\Phi_{\langle\Psi\rangle}]\!]^+ = \{\ \langle g, h\rangle \mid \exists k(\langle g, k\rangle \in [\![\Psi]\!]^+ \;\&\; \langle k, h\rangle \in [\![\Phi]\!]^+)\}$
$[\![\Phi_{\langle\Psi\rangle}]\!]^- = \{\ \langle g, g\rangle \mid \exists k\langle g, k\rangle \in [\![\Psi]\!]^+ \;\&\; \forall k(\langle g, k\rangle \in [\![\Psi]\!]^+ \Rightarrow \exists h\langle k, h\rangle \in [\![\Phi]\!]^-)\}$

We mention the following fact:

FACT 1 (Equivalences)

1. $\sim\sim\Phi$ is equivalent with Φ
2. $\Phi\,;(\Psi\,;\Upsilon)$ is equivalent with $(\Phi\,;\Psi)\,;\Upsilon$
3. $\Phi \Rightarrow [\,|\Psi \Rightarrow \Upsilon]$ is equivalent with $(\Phi\,;\Psi) \Rightarrow \Upsilon$
4. $\Phi_{\langle\pi_1\langle\pi_2\rangle\rangle}$ is equivalent with $\Phi_{\langle\pi_2;\pi_1\rangle}$

[10] A DRS is proper if it does not contain any free occurring variables, see below.

[11] We have chosen for an asymmetric treatment of disjunction, although we could have defined disjunction in a symmetric fashion (see Krahmer and Muskens (1995) and Krahmer (1995)).

For a proof of fact 1 the reader is referred to chapter 6 of Krahmer (1995).

When defining accessibility for Presuppositional DRT we have to be a little careful. In $[x \mid \textit{bathroom}(x)]\,;\,[\,\mid \textit{in strange place}(x)]$ the first occurrence of x should clearly be accessible for the occurrence of x in the condition *in strange place* (x). In $\sim[x \mid \textit{bathroom}(x)]\,;\,[\,\mid \textit{in strange place}(x)]$ this should not be the case, but in $\sim\sim[x \mid \textit{bathroom}(x)]\,;\,[\,\mid \textit{in strange place}(x)]$ accessibility should be restored again. To take care of this we shall make a distinction between *active* and *passive* discourse referents (ADR and PDR).

DEFINITION 6 (Active and Passive DRs)

1. $\mathsf{ADR}([x_1,\ldots,x_n \mid \varphi_1,\ldots,\varphi_m]) = \{x_1,\ldots,x_n\}$
   $\mathsf{PDR}([x_1,\ldots,x_n \mid \varphi_1,\ldots,\varphi_m]) = \emptyset$
2. $\mathsf{ADR}(\Phi\,;\Psi) = \mathsf{ADR}(\Phi) \cup \mathsf{ADR}(\Psi)$
   $\mathsf{PDR}(\Phi\,;\Psi) = \emptyset$
3. $\mathsf{ADR}(\sim\Phi) = \mathsf{PDR}(\Phi)$
   $\mathsf{PDR}(\sim\Phi) = \mathsf{ADR}(\Phi)$
4. $\mathsf{ADR}(\partial\Phi) = \mathsf{ADR}(\Phi)$
   $\mathsf{PDR}(\partial\Phi) = \emptyset$

Observe that $\mathsf{ADR}(\Phi_{\langle\Psi\rangle}) = \mathsf{ADR}(\Psi) \cup \mathsf{ADR}(\Phi)$, while $\mathsf{PDR}(\Phi_{\langle\Psi\rangle}) = \emptyset$. Accessibility (ACC) in some main DRS Υ is defined by setting $\mathsf{ACC}(\Upsilon) = \emptyset$ and computing the accessible discourse referents of subDRSs and subconditions in a top-down fashion.[12]

DEFINITION 7 (Accessibility)

1. If $\mathsf{ACC}(\Phi \Rightarrow \Psi) = X$, then $\mathsf{ACC}(\Phi) = X$ and $\mathsf{ACC}(\Psi) = X \cup \mathsf{ADR}(\Phi)$.
2. If $\mathsf{ACC}(\Phi \vee \Psi) = X$, then $\mathsf{ACC}(\Phi) = X$ and $\mathsf{ACC}(\Psi) = X \cup \mathsf{PDR}(\Phi)$.
3. If $\mathsf{ACC}([x_1,\ldots,x_n \mid \varphi_1,\ldots,\varphi_m]) = X$, then $\mathsf{ACC}(\varphi_i) = X \cup \{x_1,\ldots,x_n\}$. $(1 \leq i \leq m)$
4. If $\mathsf{ACC}(\Phi\,;\Psi) = X$, then $\mathsf{ACC}(\Phi) = X$ and $\mathsf{ACC}(\Psi) = X \cup \mathsf{ADR}(\Phi)$.
5. If $\mathsf{ACC}(\sim\Phi) = X$, then $\mathsf{ACC}(\Phi) = X$.
6. If $\mathsf{ACC}(\partial\Phi) = X$, then $\mathsf{ACC}(\Phi) = X$.
7. If $\mathsf{ACC}(\Phi_{\langle\Psi\rangle}) = X$, then $\mathsf{ACC}(\Psi) = X$ and $\mathsf{ACC}(\Phi) = X \cup \mathsf{ADR}(\Psi)$.

In Krahmer & Muskens (1995) and Krahmer (1995) the reader can find some discussion and justification for this way of defining accessibility, in particular where the strong connections with Karttunen (1974)'s conditions of local contexts is concerned. An occurrence of x in an atomic condition φ in Φ is said to be *free* in Φ iff $x \notin \mathsf{ACC}(\varphi)$. An occurrence of x in an atomic condition φ in a condition ψ is free in ψ iff it is *free* in $[\,\mid \psi]$. A DRS is *proper* if it does not contain any free discourse referents. Before we move on to the applications, note that the following handy lemma's hold.[13]

FACT 2 (Merging Lemma)
$[\vec{x} \mid \vec{\varphi}]\,;[\vec{y} \mid \vec{\gamma}]$ is equivalent with $[\vec{x},\vec{y} \mid \vec{\varphi},\vec{\gamma}]$,
provided no referent in $\vec{y}$ is free in any of $\vec{\varphi}$, and $[\vec{x} \mid \vec{\varphi}]$ is presupposition free.

FACT 3 (Single Negation Lemma)
Define standard negation $\neg\Phi$ as $\Phi \Rightarrow \mathbf{false}$, and let Υ be an arbitrary atomic DRS:

$\Phi\,;\sim\Upsilon$	is equivalent with	$\Phi\,;[\,\mid \neg\Upsilon]$
$\sim\Upsilon\,;\Phi$	is equivalent with	$[\,\mid \neg\Upsilon]\,;\Phi$
$\Phi \Rightarrow \sim\Upsilon$	is equivalent with	$\Phi \Rightarrow [\,\mid \neg\Upsilon]$
$\sim\Upsilon \Rightarrow \Phi$	is equivalent with	$[\,\mid \neg\Upsilon] \Rightarrow \Phi$

[12] To be precise: $\mathsf{ACC}(\Phi)$ calculates accessibility for an *occurrence* of Φ in some given DRS Φ'. Since this Φ' can be determined from the context we shall keep it hidden in the definition.

[13] In the second lemma, **false** represents the contradictory DRS which abbreviates $\sim$ **true**, with **true** as an abbreviation of $[\,\mid c \equiv c]$, for some $c \in \mathsf{CON}$.

4 APPLICATIONS

Since we slightly adapted the language of DRT we also have to revise the construction algorithm a little to make it suitable for our present purposes. The basic set-up is as follows (compare the following rule for the global structure of DRS construction with that of Kamp & Reyle (1993:86)).[14]

REVISED CONSTRUCTION ALGORITHM

Input: a discourse $S_1,\ldots,S_n$
the empty DRS $\Phi_0 = \boxed{}$
For $i = 1$ **to** n **do**:
(i) Let $\Phi_i^* = \Phi_{i-1}$; $\boxed{S_i}$. Go to (ii).
(ii) Keep on applying construction rules to each reducible condition of Φ_i^* until a DRS Φ_i is obtained that only contains irreducible conditions.

When a negation is encountered, the current DRS is prefixed with $\sim$. Moreover, every NP will be represented by a separate DRS which is immediately prefixed to the DRS in which it originated. Here is the rule for indefinites.

INDEFINITE DESCRIPTIONS RULE

Upon encountering an NP of the form 'a(n) α', replace it with a new discourse referent x and prefix the current DRS with $\boxed{\begin{array}{|c|}\hline x \\ \hline x\ \alpha \\ \hline\end{array}}$

Here '$x\ \alpha$' comes from 'a(n) α' by replacing the indefinite determiner for x. When α is an atomic predicate we shall write this as $\alpha(x)$. The rule for universally quantified NPs is modified in a similar way. As far as definite descriptions are concerned we add the following rule to the construction algorithm, which is just a presuppositional variant of the rule for indefinites.

DEFINITE DESCRIPTIONS RULE

Upon encountering an NP of the form 'the α', replace it with a new discourse referent x and prefix the current DRS with $\partial\,\boxed{\begin{array}{|c|}\hline x \\ \hline x\ \alpha \\ \hline\end{array}}$

To keep things simple pronouns are still replaced by suitable, accessible discourse referents. As an example of how this 'Revised Construction Algorithm' works, let us consider (a slight variation on) one of the notorious examples from Heim (1983).

(10) A fat man pushes his bicycle. It is broken.

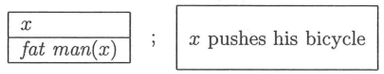
According to (10), the bicycle which went into smithereens is the one belonging to the fat man. We start with placing the first sentence in a box, thus turning it into a proto-DRS, and combining it with the initial DRS, resulting in (DRS 10).

(DRS 10) $\boxed{}$; $\boxed{\text{A fat man pushes his bicycle.}}$

[14] A more precise account would have the *syntactic analysis* of S_i as the contents of the new box. Compare Kamp & Reyle (1993).

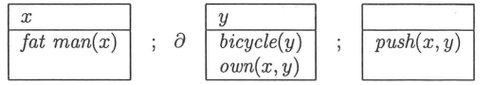
The rule for indefinite descriptions is the first which can be applied. It introduces a new discourse referent, say x, and adds the condition that x is a fat man. This gives (DRS 11), after applying the Merging Lemma to $[\,|\,]\,;[x \mid \textit{fat man}(x)]$.

(DRS 11)

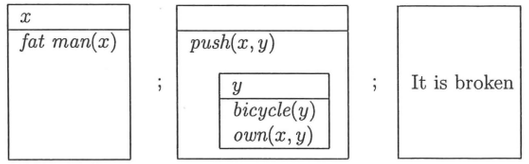


Next we come to the definite NP *his bicycle*, which we shall treat as an abbreviation of *the bicycle he owns*. This complex description introduces a new discourse referent y, while *he* picks up an accessible, suitable referent. Obviously, this will have to be x. The rest is standard DRT, and the result is (DRS 12), which may be abbreviated as (DRS 13) and extended with a new box containing the second sentence of (10).

(DRS 12)



(DRS 13)



The question is whether we can replace *it* with the intended antecedent y. We can, if y is an element of ACC(It is broken). By definition ACC(DRS 13) $= \emptyset$, so —by clause 3 of definition 7— we have to check whether y is an element of the set of active discourse referents of the representation of the first sentence (that is: ADR(DRS 12)): $\mathsf{ADR}([x \mid \textit{fat man}(x)]) \cup \mathsf{ADR}([y \mid \textit{bicycle}(y), \textit{own}(x,y)]) \cup \mathsf{ADR}([\,|\,\textit{push}(x,y)]) = \{x, y\}$. And this means that y is indeed accessible. So, what is presupposed by this DRS? According to van der Sandt it presupposes nothing: the presupposition is 'trapped'. Before we can say what the *semantic* presupposition is, we have to make a little digression.

5 Digression: calculating semantic presuppositions

An attractive feature of van Eijck's Error-state Semantics is that it is associated with an axiomatization, which allows for the calculation of the maximal presupposition of some DPL formula ϕ. For this purpose, van Eijck generalizes a method from van Eijck and de Vries (1992), in which it is shown that one can use techniques from Pratt's Dynamic Logic to reason about DPL (see also Groenendijk and Stokhof (1991:83)). Here we shall discuss a similar method for Presuppositional DRT.

The well-known semantic, Strawsonian definition of presupposing says that some sentence φ presupposes π iff whenever π is not true, φ is neither true nor false. It has been noted on many occasions that this means that the *strongest* presupposition of φ is given by the disjunction of its truth- and falsity-conditions. Thus, the strongest presupposition of a DRS Φ is given by the disjunction of the truth- and falsity-conditions of Φ. The next definition calculates truth- and falsity-conditions of arbitrary DRSs in a systematic fashion: $\mathsf{WEP}^+(\Phi, \top)$ produces a formula of Predicate Logic (PL) which is true when the DRS Φ is True, and $\mathsf{WEP}^-(\Phi, \top)$ gives a PL formula which is true when the DRS Φ is False. Using these two, we can define the semantic presupposition of a DRS Φ —$\mathsf{PR}(\Phi)$— as $\mathsf{WEP}^+(\Phi, \top) \vee \mathsf{WEP}^-(\Phi, \top)$. In the spirit of van Eijck's approach, and using the notation employed in Muskens, van

Benthem and Visser (1995) we define $\mathsf{TR}^{\pm}(\varphi)$, where φ is a condition, $\mathsf{WEP}^{\pm}(\Phi, \chi)$, where Φ is a DRS and χ a PL formula, and $\overline{\mathsf{DEF}}(\Phi)$, where Φ is again a DRS, as follows:

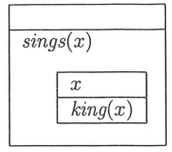
Definition 8 (WEP-calculus)

1. $\mathsf{TR}^+(\varphi) = \varphi$, for atomic φ
   $\mathsf{TR}^-(\varphi) = \neg\varphi$, for atomic φ
2. $\mathsf{TR}^+(\Phi \Rightarrow \Psi) = \neg\mathsf{WEP}^+(\Phi, \neg\mathsf{WEP}^+(\Psi, \top)) \wedge \overline{\mathsf{DEF}}(\Phi)$
   $\mathsf{TR}^-(\Phi \Rightarrow \Psi) = \mathsf{WEP}^+(\Phi, \mathsf{WEP}^-(\Psi, \top))$
3. $\mathsf{WEP}^+([\vec{x} \mid \varphi_1 \ldots \varphi_m], \chi) = \exists\vec{x}(\mathsf{TR}^+(\varphi_1) \wedge \ldots \wedge \mathsf{TR}^+(\varphi_m) \wedge \chi)$
   $\mathsf{WEP}^-([\vec{x} \mid \varphi_1 \ldots \varphi_m], \chi) = \forall\vec{x}(\mathsf{TR}^-(\varphi_1) \vee \ldots \vee \mathsf{TR}^-(\varphi_m)) \wedge \chi$
4. $\mathsf{WEP}^+(\Phi\,;\Psi, \chi) = \mathsf{WEP}^+(\Phi, \mathsf{WEP}^+(\Psi, \chi))$
   $\mathsf{WEP}^-(\Phi\,;\Psi, \chi) = \neg\mathsf{WEP}^+(\Phi, \neg\mathsf{WEP}^-(\Psi, \top)) \wedge \overline{\mathsf{DEF}}(\Phi) \wedge \chi$
5. $\mathsf{WEP}^+(\sim\Phi, \chi) = \mathsf{WEP}^-(\Phi, \chi)$
   $\mathsf{WEP}^-(\sim\Phi, \chi) = \mathsf{WEP}^+(\Phi, \chi)$
6. $\mathsf{WEP}^+(\partial\Phi, \chi) = \mathsf{WEP}^+(\Phi, \chi)$
   $\mathsf{WEP}^-(\partial\Phi, \chi) = \mathsf{WEP}^-(\mathbf{true}, \chi)$[15]
7. $\overline{\mathsf{DEF}}(\Phi) = \mathsf{WEP}^+(\Phi, \top) \vee \mathsf{WEP}^-(\Phi, \top)$, if Φ is $\partial\Psi$ or $[\vec{x} \mid \varphi_1, \ldots, \varphi_m]$
8. $\overline{\mathsf{DEF}}(\sim\Phi) = \overline{\mathsf{DEF}}(\Phi)$
9. $\overline{\mathsf{DEF}}(\Phi\,;\Psi) = \overline{\mathsf{DEF}}(\Phi) \wedge \neg\mathsf{WEP}^+(\Phi, \neg\overline{\mathsf{DEF}}(\Psi))$

To illustrate this calculus let us first work out a simple example. Consider (11) and its representation (DRS 14), which abbreviates $\partial[x \mid \textit{king}(x)]\,;[\,|\,\textit{sings}(x)]$.

(11) The king sings.

(DRS 14)



We first calculate $\mathsf{WEP}^+((\text{DRS } 14), \top)$:

$$\begin{array}{l}\mathsf{WEP}^+(\partial[x \mid \textit{king}(x)]\,;[\,|\,\textit{sings}(x)], \top) \Leftrightarrow\\ \mathsf{WEP}^+(\partial[x \mid \textit{king}(x)], \mathsf{WEP}^+([\,|\,\textit{sings}(x)], \top)) \Leftrightarrow\\ \mathsf{WEP}^+([x \mid \textit{king}(x)], \mathsf{WEP}^+([\,|\,\textit{sings}(x)], \top)) \Leftrightarrow\\ \exists x(\mathsf{TR}^+(\textit{king}(x)) \wedge \mathsf{WEP}^+([\,|\,\textit{sings}(x)], \top)) \Leftrightarrow\\ \exists x(\textit{king}(x) \wedge \textit{sings}(x))\end{array}$$

In words: (DRS 14) is True if there is a king who sings. Next, let us see when it is False. For this purpose, we determine $\mathsf{WEP}^-((\text{DRS } 14), \top)$:

$$\begin{array}{l}\mathsf{WEP}^-(\partial[x \mid \textit{king}(x)]\,;[\,|\,\textit{sings}(x)], \top) \Leftrightarrow\\ \forall x(\textit{king}(x) \rightarrow \neg\textit{sings}(x)) \wedge \exists x(\textit{king}(x))\end{array}$$

[15] The intuition here is that the tautological DRS **true** (cf. footnote 13) is rejected whenever the presuppositional DRS $\partial\Phi$ is: never.

In words: (DRS 14) is False if there is a king, but no king sings. The disjunction of $\mathsf{WEP}^+((\text{DRS } 14), \top)$ and $\mathsf{WEP}^-((\text{DRS } 14), \top)$ gives the presupposition of (DRS 14), and it is easily seen that this presupposition is that there is a king. In general, the following fact can be proven about these five functions.[16]

FACT 4 (From Presuppositional DRT to PL) For all models M and assignments g:

1. $g \in [\![\mathsf{TR}^+(\varphi)]\!]^{PL}_M \Leftrightarrow g \in [\![\varphi]\!]^+_M$
   $g \in [\![\mathsf{TR}^-(\varphi)]\!]^{PL}_M \Leftrightarrow g \in [\![\varphi]\!]^-_M$
2. $g \in [\![\mathsf{WEP}^+(\Phi, \chi)]\!]^{PL}_M \Leftrightarrow \exists h(\langle g, h\rangle \in [\![\Phi]\!]^+ \;\&\; h \in [\![\chi]\!]^{PL}_M)$
   $g \in [\![\mathsf{WEP}^-(\Phi, \chi)]\!]^{PL}_M \Leftrightarrow \exists h(\langle g, h\rangle \in [\![\Phi]\!]^- \;\&\; h \in [\![\chi]\!]^{PL}_M)$
3. $g \in [\![\overline{\mathsf{DEF}}(\Phi)]\!]^{PL}_M \Leftrightarrow \mathrm{DEF}_{M,g}(\Phi)$

Where $[\![\varphi]\!]^{PL}_M$ is the Groenendijk & Stokhof formulation of the Tarskian semantics ($\models_t$) of Predicate Logic (see Groenendijk & Stokhof (1991:72)). Thus $[\![\varphi]\!]^{PL}_M = \{g \mid M, g \models_t \varphi\}$. For a proof of fact 4 the reader is referred to Krahmer (1995). Given that $\mathsf{PR}(\Phi)$ abbreviates $\mathsf{WEP}^+(\Phi, \top) \vee \mathsf{WEP}^-(\Phi, \top)$, it is easily seen that $\mathsf{PR}(\Phi)$ is indeed the strongest presupposition of Φ. When we assume that Φ triggers no presuppositions, then $\mathsf{PR}(\Phi_{\langle\Psi\rangle}) \Leftrightarrow \mathsf{WEP}^+(\Psi, \top)$. In other words, $\Phi_{\langle\Psi\rangle}$ presupposes the Truth of Ψ. Here is PR(DRS 12).[17]

$$\exists x(\mathit{fat}\ \mathit{man}(x) \wedge \exists y(\mathit{bike\text{-}of}(y,x) \wedge \mathit{push}(x,y))) \vee$$
$$\forall x(\mathit{fat}\ \mathit{man}(x) \rightarrow \exists y(\mathit{bike\text{-}of}(y,x) \wedge \forall y(\mathit{bike\text{-}of}(y,x) \rightarrow \neg \mathit{push}(x,y))))$$

In words: either there is a fat man who has a bicycle which he pushes or every fat man doesn't push the bike(s) he owns. Thus, on the semantic side Presuppositional DRT does not give rise to the Karttunen and Peters (1979) binding-problem when presupposition and quantification interact, and neither does it predict Heim (1983)'s overly strong presuppositions in these cases. In this respect Presuppositional DRT makes predictions which are closely related to those of Beaver's Kinematic Predicate Logic and, in particular, Van Eijck's Error-state semantics.

6 AGAIN: PRESUPPOSITIONS AS ANAPHORA

But above we noted that Presuppositional DRT is also perfectly compatible with van der Sandt's presuppositions-as-anaphora approach. We also claimed that van der Sandt's theory could benefit from defining the resolution algorithm on top of Presuppositional DRT instead of ordinary DRT. To back-up this claim, let us re-discuss the advantages briefly.

First, Presuppositional DRT indeed allows for the interpretation of presuppositional DRSs. This entails that the constraints on accommodation (which refer to semantic concepts such as satisfaction and consistency) can be applied on the spot, even though other presuppositional DRSs are still unresolved. In this way, the procedural version of the resolution algorithm can stop when it has found a

[16] This fact relates PL with the version of Presuppositional DRT using total assignments (which is obtained from definition 2 by replacing the set of finite assignments F for the set of total assignments G, and replacing $g[\vec{x}]h$ for $g[\vec{x}]h$, which abbreviates $\forall y(y \notin \{\vec{x}\} \Rightarrow g(y) = h(y))$). It should be stressed that this is completely harmless since the Revised Construction Algorithm still only produces proper DRSs, while the switch from finite to total assignments only has a semantic effect in the case of non-proper DRSs.

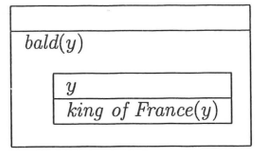
[17] With $\mathit{bike\text{-}of}(y,x)$ as shorthand for $(\mathit{bike}(y) \wedge \mathit{own}(x,y))$.

solution, which by the very nature of the method sketched in section 2 will be the most preferred one.[18]

Another issue was the accommodation of failing presuppositions. Re-consider example (5), repeated here as (13), and its representation generated by the present Revised Construction Algorithm, (DRS 15).

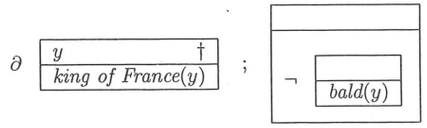
(13) It is not the case that the king of France is bald.

(DRS 15)



Van der Sandt's resolution-algorithm predicts that accommodating the presuppositional DRS as high as possible is the preferred resolution. However, as discussed, the presuppositional DRS loses its 'presuppositionhood' in the process, hence the outcome will be false in any model where the French have no king. Yet intuitively, as Kracht (1994) puts it, an accommodated presupposition should still be a presupposition. In Presuppositional DRT this is an easy matter: we accommodate the presuppositional DRS *as a presuppositional DRS* and not as an ordinary DRS. Schematically, global accommodation of Ψ in $\sim(\Phi_{\langle\Psi\rangle})$ results in $(\sim\Phi)_{\langle\Psi\rangle}$ (recall that this is just an abbreviation of $\partial\Psi\,;\sim\Phi$), and not in $\Psi\,;\sim\Phi$. Modulo the Single Negation lemma, this means that the (slightly modified) presupposition resolution algorithm delivers (DRS 16) when applied to (DRS 15) (where † again indicates that the presuppositional DRS has been resolved).

(DRS 16)



It is readily seen that (DRS 16) presupposes the existence of a French monarch (just calculate PR(DRS 16)), so this DRS will indeed be undefined when the presupposition that there is a king of France is not satisfied.[19]

[18] It should be noted that a certain amount of backtracking is inevitable, in particular when a sentence contains two or more interdependent presuppositions. A case in point is the following example:

(12) Either the king of France opened the exhibition or the president of France opened it.

Suppose the left disjunct is resolved first; global accommodation of the presupposition that France has a king will not be blocked, although evaluation of the right disjunct will show that the presupposition should be *locally* accommodated. Of course, this kind of backtracking can be avoided when actually implementing Van der Sandt's resolution algorithm, for instance using a rule of thumb which first determines which presuppositions are interdependent.

[19] But what about the second reading of (13)? Consider:

(14) It is not the case that the king of France is bald, since there is *no* king of France.

This example does not presuppose the existence of a king of France. Van der Sandt accounts for this reading by *locally* accommodating the presupposition. If we locally accommodate the presupposition *as a presupposition*, the resulting DRS will end up undefined, which is plainly wrong. One possibility is to accommodate presuppositions only as presuppositions in the main DRS. A more principled possibility is the following: we *always* accommodate a presupposition as a presupposition, but in certain specially marked contexts (such as *echo-negations* like (14), cf. Sandt (ms.)) the presupposition operator is 'neutralized' by a *presupposition wipe-out device*, going back to Bochvar (1939)'s *meta-assertion* operator. For more details: Beaver & Krahmer (1995).

An advantage of the fact that Presuppositional DRT is based on Double Negation DRT has to do with the treatment of disjunction. In standard DRT (the footing of van der Sandt's theory) it is predicted that the left disjunct is not accessible for anaphoric elements in the right disjunct. In Double Negation DRT things are different: the *negation* of the left disjunct is accessible for the right disjunct. Re-consider the relevant examples (6) and (7). Both examples presuppose nothing. Intuitively, both examples should be rendered in terms of binding and this is exactly what happens when van der Sandt's resolution mechanism is applied to Presuppositional DRT. Here is the schematic representation of the two examples under discussion.

(DRS 17)

$\sim\Phi \vee \Psi_{\langle\Upsilon\rangle}$

Where Φ represents 'there is a bathroom in this house/Mary's autobiography has appeared', Υ represents the elementary presupposition 'there is a bathroom/Mary has had a book published' and Ψ 'it is in a strange place/John must be very proud of it'. Given the calculation of accessibility, it is easily seen that the passive discourse referents of $\sim\Phi$ are accessible for the presuppositional DRS Υ. And since the passive discourse referents of $\sim\Phi$ are the *active* discourse referents of Φ, both the bathroom and the autobiography are accessible for the presuppositions in Υ: binding is possible and preferred in both cases.

Finally, as far as presupposition-quantification interaction is concerned, let me note that in Presuppositional DRT it is no longer predicted that *Every German loves his car* is equivalent to *Every German who has a car loves it.* The presupposition triggered by *his car* is still accommodated in the scope of *every*, but in the process remains presuppositional, and this destroys the equivalence.

7 Discussion: the meaning of resolution

Given that we can interpret DRSs before *and* after resolution, a natural question to ask is whether resolving presuppositions preserves meaning. Notice that this also tells us something about the relationship between the two approaches to presupposition compatible with Presuppositional DRT; the semantic and the algorithmic one. Every time a resolution preserves meaning, the semantic and the algorithmic approach make the same predictions. After all: take some DRS Φ. The semantic presupposition of Φ is given by $\mathsf{PR}(\Phi)$. Feeding Φ to van der Sandt's resolution algorithm produces a DRS Φ'. If Φ and Φ' are equivalent (thus: the resolution preserves meaning), then $\mathsf{PR}(\Phi')$ must be equivalent to $\mathsf{PR}(\Phi)$ as well. So, when the resolution preserves meaning, this tells us something about the *similarities* between the two approaches. And when it does not, this tells us something about the *differences* between the two.

It will not surprise the reader that in general binding can be understood in a meaning preserving way,[20] but that accommodation is an entirely different kettle of fish. For instance, (DRS 4) and (DRS 5) are obviously not equivalent. Since accommodation is not a meaning preserving operation in general, differences between the two approaches arise. This is one of the major points where van der Sandt's theory positively distinguishes itself from other, purely semantic approaches to presuppositions, which predict *weaker* presuppositions in these cases.

Why is accommodation not meaning-preserving? In one sense this is just what we would expect. After all, Lewis (1979) describes accommodation as the phenomenon that missing information simply 'springs into existence': the representation of the foregoing is altered in such a way that the presupposition is satisfied after all. And it is no surprise that this changes the meaning. On the other hand,

[20] See Krahmer (1995) for discussion.

consider (DRS 15), which is of the form $\sim(\Phi_{\langle\Psi\rangle})$. If we feed this DRS to the (slightly adapted) resolution algorithm, the presuppositional DRS will be accommodated as a presuppositional DRS and the result is (DRS 16), which is of the form $(\sim\Phi)_{\langle\Psi\rangle}$ These two DRSs differ in their *dynamic* meaning (thus, they are not equivalent), and hence even this form of accommodation is not entirely meaning-preserving. Yet, we could reasonably expect the accommodation to preserve meaning in this case; note for instance that $\mathsf{PR}(\sim(\Phi_{\langle\Psi\rangle}))$ is equivalent with $\mathsf{PR}((\sim\Phi)_{\langle\Psi\rangle})$. In fact, this points to a weakness in the *interpretation* of Presuppositional DRT.

What is the problem here? Recall that in dynamic semantics, a single negation is understood as a plug with respect to anaphoric reference: it blocks dynamic potential. Therefore, *no* discourse referent which originates in the scope of a single negation can be referred back to. So, in $\sim(\Phi_{\langle\Psi\rangle})$ any referent introduced in either Φ or Ψ is inaccessible outside the scope of the negation. In $(\sim\Phi)_{\langle\Psi\rangle}$ things are different; referents introduced in the main universe of Ψ *are* available for future reference. One would expect presuppositions to be totally insensitive to negation, and that this is not the case for the interpretation of Presuppositional DRT given in definition 2 (nor for the related interpretations of e.g. Kinematic Predicate Logic and Error-state Semantics for DPL) is an artefact of the dynamic interpretation of single negations. What we want is the following: a single negation should be a *plug* for anaphoric reference to non-presupposed discourse referents (like those representing indefinites), but at the same time it should be a *hole* for anaphoric reference to presupposed discourse referents (like those introduced in the representation of definite descriptions). Compare:[21]

(15) a. John didn't bake the pie for Mary's birthday. It was rather tasty.

b. John didn't bake a pie for Mary's birthday. ?? It was rather tasty.

The phenomenon of a logical connective which should act as both a plug *and* a hole with respect to anaphoric reference is by no means restricted to negation. Consider:

(16) a. If an inhabitant of France is bald, he will wear a curled wig. ?? He is a dedicated follower of fashion.

b. If the king of France is bald, he will wear a curled wig. He is a dedicated follower of fashion.

Again, we want to block anaphoric reference to an indefinite NP in the antecedent of a conditional, but not to a definite one in the same position. Can we do something about this?

Intuitively a presupposition is something which should be the case beforehand. In what sense, does $\Phi_{\langle\Psi\rangle}$ conform to this intuition? If we evaluate $\Phi_{\langle\Psi\rangle}$ with respect to an assignment g, then the presuppositional DRS Ψ is supported if there is an assignment h such that $\langle g,h\rangle \in [\![\Psi]\!]^+$; presupposition is a matter of *successors*. Now consider the following alternative definition:[22]

Definition 9 ($\Phi_{\langle\Psi\rangle}$ (re-)defined)
$[\![\Phi_{\langle\Psi\rangle}]\!]^+ = \{\langle g,h\rangle \mid \exists k\langle k,g\rangle \in [\![\Psi]\!]^+ \;\&\; \langle g,h\rangle \in [\![\Phi]\!]^+\}$
$[\![\Phi_{\langle\Psi\rangle}]\!]^- = \{\langle g,h\rangle \mid \exists k\langle k,g\rangle \in [\![\Psi]\!]^+ \;\&\; \langle g,h\rangle \in [\![\Phi]\!]^-\}$

Arguably this definition comes closer to the intuition of a presupposition as something which holds already. Definition 9 supports the following fact, as the reader can easily verify.

[21] In (15.b) the indefinite is to be understood in a non-specific way.

[22] Notice that this interpretation of $\Phi_{\langle\Psi\rangle}$ is not definable in terms of sequencing and *any* alternative interpretation of ∂, since $[\![\partial\Psi\,;\,\Phi]\!]^-$ will always be a subset of the diagonal of F^2 $(= \{\langle g,h\rangle \mid g = h\})$.

FACT 5 (Equivalences)
$\sim(\Phi_{\langle\Psi\rangle})$ is equivalent with $(\sim\Phi)_{\langle\Psi\rangle}$
$[\,|\,\Phi_{\langle\Psi\rangle} \Rightarrow \Upsilon]$ is equivalent with $[\,|\,\Phi \Rightarrow \Upsilon]_{\langle\Psi\rangle}$

This means that if we interpret presuppositional DRSs as in definition 9, the resolution of (DRS 15) yielding (DRS 16) *does* preserve meaning. Moreover, the treatment of (15) *and* (16) has become possible for the semantic account as well.

To accommodate the 'looking-back' character of the new presupposition operator two additional modifications are needed. First of all, it does not make much sense any more to define the truth-combinations with respect to Λ (the empty-assignment). After all, you cannot 'look back' from the empty assignment. Therefore the truth-combinations are re-defined with respect to an arbitrary assignment g (compare the DPL notion of truth). For similar reasons, we have to generalize the notion of 'domain-extension', i.e. $g\{\vec{x}\}h$ now comes to mean 'assignment h extends assignments g *at least* with $\vec{x}$ '. Formally: $g\{\vec{x}\}h$ abbreviates $\mathsf{DOM}(g) \cup \{\vec{x}\} \subseteq \mathsf{DOM}(h)$ & $\forall y \in \mathsf{DOM}(g) : g(y) = h(y)$. Otherwise, the interpretation of Presuppositional DRT remains as it was. For more details the reader is referred to Krahmer (1995).

8 CONCLUDING REMARKS

In this paper we have discussed Presuppositional DRT. The system of Presuppositional DRT can be seen as a combination of Krahmer & Muskens' Double Negation DRT with the static partial approach to presuppositions of Peters (1975).[23] We have seen that presuppositions can be studied using Presuppositional DRT in two, rather different ways. Given some DRS Φ we can either calculate $\mathsf{PR}(\Phi)$ and find the semantic presupposition, or we can feed Φ to van der Sandt's resolution algorithm. It was shown that the resolution algorithm actually benefits from the partial interpretation of Presuppositional DRT, even though the algorithm itself remains essentially as van der Sandt developed it. In the previous section it was shown that this picture also allows us to study the differences in prediction between the two approaches. In connection with this we discussed an alternative interpretation for Presuppositional DRT. Although it needs to be fleshed out in more detail, it seems to be a promising line for future research. The alternative definition makes it possible to bridge the 'accommodation-gap' between the two approaches to presupposition discussed in this paper.

In fact, for *both* approaches a better understanding of accommodation seems desirable. From the semantic perspective accommodation is a means to avoid presuppositions which are too weak. It is easily seen that the new interpretation discussed in the previous section does not lead to a strengthening of the predicted semantic presuppositions. After all, its only novelty with respect to the original interpretation concerns the dynamic potential of discourse referents introduced in presuppositional DRSs, and the semantic notion of presupposition is merely defined in terms of truth and falsity. So as things stand, the semantic approach still predicts presuppositions which are too weak in certain cases.[24] On the other hand, van der Sandt's approach to accommodation has a somewhat 'syntactic' flavor: it works very well, but it is not entirely clear *why* it works so well. An interesting alternative might be to look at van der Sandt's resolution-algorithm from a *proof-theoretic* perspective. This might go as follows: in Ahn and Kolb (1990) a function

[23] In Krahmer (1995) the relation between static and dynamic approaches to presupposition is discussed at length. One of the conclusions is that the notorious examples involving quantifier-presupposition interaction from Karttunen and Peters (1979) and Heim (1983) can be dealt with inside an ordinary partial interpretation of *static* Predicate Logic. Dynamics is not needed to deal with them.

[24] But see for instance Beaver (1993) for a defense of weak presuppositions.

τ is defined which maps DRSs to expression of *Constructive Type Theory* (CTT). The resulting CTT expressions $\tau(\Phi)$ are evaluated given a context/set of premisses Γ. Now suppose we re-define τ to map *Presuppositional* DRT to CTT, and that we evaluate $\tau(\Phi_{\langle\Psi\rangle})$ in some context Γ. Then three things can happen: (i) the presupposition follows from Γ, it is 'bound'. This would be the preferred situation. (ii) the presupposition does not follow from Γ. Then either we have missed something; our context Γ is incomplete and should be extended ('global accommodation'), or, we want to leave our set of premisses untouched and modify $\tau(\Phi_{\langle\Psi\rangle})$ ('local accommodation'). (iii) Γ *contradicts* the presupposition. What to do? The best option is to give up and ask the speaker for clarification. This all to brief exposé indicates that it might be interesting to look at van der Sandt's resolution algorithm from a proof-theoretic point of view.

ACKNOWLEDGEMENTS Thanks are first and foremost due to Reinhard Muskens who commented on several previous versions of this paper, and made many highly useful suggestions. Thanks are also due to David Beaver, Jan van Eijck, Bart Geurts, Jan Jaspars, Rob van der Sandt, Elias Thijsse and Kees Vermeulen for various kinds of support. Finally I want to thank the audience of the Tenth Amsterdam Colloquium, as well as the audiences in ParADySe (Utrecht, 1995) and at the IMS (Stuttgart, 1995), for their remarks and suggestions. Needless to say, I alone should be held responsible for any errors.

REFERENCES

Ahn, R. and Kolb, H.-P.: Discourse representation meets constructive mathematics, in L. Kálmán and L. Pólós (eds.), *Papers from the 2nd Symposium on Logic and Language*, Akadémia Kiadó, Budapest, 1990.

Beaver, D.: 1992, The kinematics of presupposition, in P. Dekker and M. Stokhof (eds.), *Proceedings of the Eight Amsterdam Colloquium*, ILLC, Amsterdam.

Beaver, D.: 1993, *What Comes First in Dynamic Semantics*, ILLC Report LP-93-15, Amsterdam.

Beaver, D.: 1994, When variables don't vary enough, in M. Harvey and L. Santelman (eds.), *Proceedings of SALT IV*, Cornell University, Ithaca, 35-60.

Beaver, D.: 1995, *Presupposition and Assertion in Dynamic Semantics*, Ph.D. thesis, Edinburgh.

Beaver, D. and Krahmer, E.: 1995, *Back to the Future*, Unpublished manuscript.

Berg, M.v.d.: 1991, Dynamic generalized quantifiers, in J.v.d. Does and J.v. Eijck (eds.), *Generalized Quantifiers: Theory and Applications*, to appear as CSLI Lecture Notes, Stanford.

Berg, M.v.d.: 1993, Full dynamic predicate logic, in K. Bimbó and A. Máté (eds.), *Proceedings of the 4th Symposium on Logic and Language*, Budapest.

Blamey, S.: 1986, Partial logic, in D. Gabbay and F. Guenthner (eds.) *Handbook of Philosophical Logic, Vol. 3*, Reidel, Dordrecht, 1-70.

Bochvar, D.: 1981 (1939), On a three-valued logical calculus and its applications in the analysis of paradoxes of the classical extended functional calculus, *History and Philosophy of Logic* 2, 87-112.

Dekker, P.: 1993, *Transsentential Meditations: Ups and Downs in Dynamic Semantics*, Ph.D. thesis, ILLC, Amsterdam.

Eijck, J.v.: 1993a, The dynamics of description, *Journal of Semantics* 10, 239-267.

Eijck, J.v.: 1993b, *Presuppositions and Dynamic Logic*, Manuscript, CWI, to appear in M. Kanazawa et al. (eds.), *Proceedings from the 2nd CSLI Workshop on Logic, Language and Computation*, CSLI, Stanford.

Eijck, J.v.: 1994, Presupposition failure - a comedy of errors, in *Formal Aspects of Computing* 6A, 766-787.

Eijck, J.v. and Vries, F.-J.d.: 1992, Dynamic interpretation and Hoare deduction,

in *Journal of Language, Logic and Information*, 1, 1-44.
Fernando, T.: 1994, Translations in the dynamic analysis of presuppositions, to appear in *Proceedings of the 5th Symposium on Logic and Language*.
Geurts, B.: 1995, *Presupposing*, Ph.D. thesis, Osnabrück
Groenendijk, J. and Stokhof, M.: 1991: Dynamic Predicate Logic, *Linguistics and Philosophy* 14, 39-100.
Heim, I.: 1983, On the projection problem for presuppositions, in M. Barlow (ed.), *Proceedings of WCCFL2*, Stanford University, Stanford, 114-125.
Kamp, H. and Reyle, U.: 1993, *From Discourse to Logic*, Kluwer, Dordrecht.
Karttunen, L.: 1974, Presupposition and linguistic context, *Theoretical Linguistics* 1, 181-194.
Karttunen, L. and Peters, S.: 1979, Conventional implicature, in C. Oh and D. Dinneen (eds.), *Syntax and Semantics 11* Academic Press, New York.
Kracht, M.: 1994, Logic and control: how they determine the behaviour of presuppositions, in J.v. Eijck and A. Visser (eds.) *Logic and Information Flow*, MIT Press, Cambridge, Mass.
Krahmer, E.: 1994, Partiality and dynamics, in P. Dekker and M. Stokhof (eds.), *Proceedings of the Ninth Amsterdam Colloquium*, ILLC, Amsterdam.
Krahmer, E.: 1995, *Discourse and Presupposition*, Ph.D. thesis, Tilburg, available from `http://tkiwww.kub.nl:2080/tki/Faces/Ek/ek.html`.
Krahmer, E. and Muskens, R.: 1994, Of umbrella's and bathrooms, in M. Harvey and L. Santelman (eds.), *Proceedings of SALT IV*, Cornell University, Ithaca.
Krahmer, E. and Muskens, R.: 1995, Negation and disjunction in DRT, *Journal of Semantics* 12, 357-376.
Lewis, D.: 1979, Scorekeeping in a language game, *Journal of Philosophical Logic* 8, 339-359.
Muskens, R., Benthem, J.v. and Visser, A.: 1995, Dynamics, in J.v. Benthem and A.t. Meulen (eds.), *Handbook of Logic in Linguistics*, to appear with Elsevier.
Peters, S.: 1975, A truth-conditional formulation of Karttunen's account of presupposition, in *Texas Linguistic Forum*, Austin, 137-149.
Piwek, P.: 1993, Partiële Discourse Representatie Theorie voor Neutrale Perceptieverslagen, Master's thesis, Tilburg.
Sandt, R.v.d.: 1989, Presupposition and discourse structure, in R. Bartsch et al. (eds.), *Semantics and Contextual Expression*, Foris, Dordrecht.
Sandt, R.v.d.: 1992, Presupposition projection as anaphora resolution, *Journal of Semantics* 9, 333-377.
Sandt, R.v.d.: Ms, Discourse semantics and echo-quotation, to appear in *Linguistics and Philosophy*.
Sandt, R.v.d. and Geurts, B.: 1991, Presupposition, anaphora and lexical content, in O. Herzog et al. (eds.) *Text Understanding in LILOG*, Springer, Berlin.
Zeevat, H.: 1993, Presupposition and accommodation in update semantics, *Journal of Semantics* 9, 379-412.

Integrating Situations into a Theory of Discourse Anaphora

David Milward
Universität des Saarlandes

Abstract[1]

This paper provides an account of definite and pronominal anaphora which draws together insights from dynamic semantics and situation semantics. The result is a dynamic semantics in which contexts consist of the parts of a scenario which have been brought to a hearer's attention. The contexts are structured by meta-level disjunction, corresponding to a set of situations where one is in attention but the hearer doesn't know which one, and meta-level conjunction corresponding to a set of situations which are simultaneously in attention. Data from *bridging* reference, disjunction, and *telescoping* is used to motivate the account, together with some of the data which has been used in the past to argue for and against situation based accounts of anaphora.

1 Theories of Discourse Anaphora

It is useful to separate out three treatments of anaphora. The first, and most traditional approach is to try to fill in the anaphor, either at the level of word string, syntax, or semantic representation. For pronominal anaphora almost all recent proposals have been at the level of semantic representation. Perhaps the most sophisticated approach is the E-type approach (original work by Evans 1980, but see e.g. Does 1994 for recent developments). Consider the following discourse.

(1) A young child came in today.
He sat down at the computer and started typing.

The E-type approach makes the second sentence less context dependent by replacing *He* by a representation equivalent to *The young child who came in today.*

There has been much valuable, but perhaps inconclusive, discussion of whether an E-type approach, in treating pronouns similar to definites, requires that the description must refer uniquely and whether or not that is problematic (e.g. Heim 1990, Kadmon 1990, Does 1993, Dekker 1993). A rather different issue concerns the treatment of mutually dependent anaphora. Consider the discourses:

(2) a. A colleague of mine looked after a baby last week.
After 5 minutes **it** had driven **him** crazy.
b. A girlfriend of a colleague of mine is a journalist.
She tells **him** all the latest Royal gossip.

In (a), **it** refers to a baby looked after by a colleague, **him** to a colleague who looked after *that* baby. In E-type approaches we get an interpretation along the lines of:

(3) After 5 minutes the baby who was looked after by a colleague of mine last week drove the colleague of mine who looked after a baby last week crazy.

For (b) we would need an interpretation along the lines of:

(4) The girlfriend of a colleague of mine who is a journalist tells the colleague of mine who has a girlfriend who is a journalist all the latest Royal gossip.

These interpretations are certainly inelegant, but this might not be seen as a real problem for E-type analyses provided we can formulate the appropriate rules to create them from the surface syntax. However, there is certainly a problem when we replace indefinites by disjunctions. Consider e.g.

1. I am indebted to Paul Dekker, Claire Gardent, Jonathan Ginzburg and Roger Kibble for comments on earlier drafts, and to the European Community for funding as part of the FraCaS project, LRE 62-051

(5) a. Mary can take the truck, or Sue can take the pickup. **She** can park **it** by the entrance.
b. We can get a young girl to recite a poem, or an older one to read a passage in French. If **it** takes **her** about 10 minutes then ...

Here we again have mutually dependent anaphors, but it is much less clear how to construct an appropriate description to replace the anaphors.

Instead of attempting to make sentences context independent, the second approach to anaphora builds semantics only for discourses as a whole, not for the individual sentences[2]. A simple approach which has been used in the Artificial Intelligence tradition going back to e.g. Winograd 1976 is to treat indefinites as setting up new 'knowledge based entities'. Subsequent anaphors then add further properties to this entity. For example, after processing

(6) A man sleeps

we set up a new index, i1, and assert of that index that it is a man and it sleeps. On encountering:

(7) He snores

we merely add the information that index i1 snores. Now consider processing the sentence:

(8) A man buys a picture

Here two indices would be created, e.g. i1 and i2. Taking an object orientated approach to knowledge representation, we have a choice of putting the information under either i1 or i2, or under both. Putting it under both gives:

i1
man(i1)
buys(i1,i2)

i2
picture(i2)
buys(i1,i2)

The resulting representations are similar to the file cards of Heim 1982.

An alternative to an object orientated approach is to assert information concerning tuples of knowledge based entities. This can be represented as follows:

i1 i2
man(i1)
picture(i2)
buys(i1,i2)

This representation is identical to the DRS structure which would be built for the same sentence. In fact we can think of DRT as a better worked out version of the early AI approaches, with the further advantage of having a well defined semantics for its representations. Similar to the AI approaches, DRT achieves semantic binding for inter-sentential anaphora by adding to an existing structure. This is achieved by the DRS construction algorithm, which we can think of as an operation upon representations, or, more controversially, as an operation on types (this perspective is given by the situation theoretic reformulation of DRT, Cooper 1993b). For intra-sentential anaphora, e.g. *donkey* anaphora, DRT is compatible with two approaches. The first is where binding is achieved not at the level of representation, but at the level of interpretation. For example, the rule for implication in Kamp and Reyle 1993 makes the implication true if all assignments (c.f. individual cases, instantiations or embeddings) which satisfy the antecedent can be extended to assignments which satisfy the consequent. However, Kamp and Reyle (p149) emphasise that this is only due to some technical aspects of the construction algorithm. The rules could be reformulated so that binding is done at the level of the types, with the antecendent

2. Of course we can regain a notion of sentence meaning as a mapping between contexts, the essential notion behind dynamic semantics. We will return to this.

DRS copied explicitly into the consequent. This would then allow a purely static interpretation of DRSs.

The third approach to anaphora is where we consider each sentence to be evaluated in context, and to map to a new context. This is the essential characteristic of dynamic semantics approaches e.g. Barwise 1987, Groenendijk and Stokhof 1991. This approach differs from the first kind since there is no attempt to make the sentences themselves context independent. It also differs from the second kind of approach in being more general. Any example of the second kind, such as DRT, can be given a dynamic reinterpretation, since we can regard sentence meanings as mappings from the meaning of the discourse before the sentence was processed, to the meaning of the discourse afterwards (e.g. as a mapping between two DRSs). For example, the sentence *a man buys a picture* evaluated in the empty discourse context gives the mapping:

□ [[a man buys a picture]]

x y
man(x)
picture(y)
buys(x,y)

Standard DRT is a special case of dynamic interpretation, since the contexts are identified with the meaning of the discourse upto that point. This conflation of the meaning of the discourse and the context has important consequences which we will examine in the next section. For now, consider a non-standard DRT where DRSs represent contexts, but not discourse meanings (see Fernando 1994 for discussion along these lines). In such a system it is natural to interpret a DRS as a function from values of referents to truth values (i.e. as a relation). The DRS can then be replaced by a lambda expression such as:

λx.λy. man(x) & picture(y) & buys(x,y)

At first sight this replacement seems surprising. An essential feature of DRT is the treatment of discourse variables as distinct from bound variables. How can a lambda expression capture non local binding? The trick is to bind to argument *positions* rather than to variables.[3] Taking this perspective allows a unifying view of accounts such as Predicate Logic with Anaphora, Dekker 1994 and STDRT, Cooper 1993b.

In Predicate Logic with Anaphora (PLA) contexts are sets of tuples of model theoretic individuals. This is equivalent to evaluating a function or relation in a particular model. For example, if a model supports John buying picture 1, and Fred buying picture 2, the lambda expression above would evaluate to the following set of tuples:

{ ⟨ john, picture1 ⟩,
⟨ fred, picture2 ⟩ }

john and *fred* share the same *column* i.e. they are in the same position in their respective tuples. *john* and *picture1* share the same *row* i.e. are members of the same tuple. Dekker's approach to binding is to associate pronouns with particular columns, which correspond to particular argument positions of a lambda expression

As an alternative to using tuples, where items have a particular ordering, we can extract the ordering information and get a set of sets of pairs i.e.

{ {(1,john),(2,picture1)},
{(1,fred),(2,picture2)} }

However, the choice of names 1 and 2 is arbitrary. We could just have well used the names of the original discourse referents i.e.

{ {(x,john),(y,picture1)},
{(x,fred),(y,picture2)} }

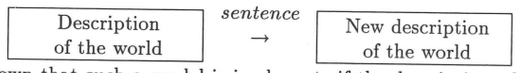
3. The binding process is trivial if a pronoun picks an argument position from the current context (this is similar to picking a particular column in PLA, described in the next paragraph). Things are more complicated if binding is done via syntactic coindexing.

The result is a set of partial assignments. Partial or total assignments have been the basis of most of the dynamic semantics proposed so far e.g. Barwise 1987, Groenendijk and Stokhof 1991.

In the STDRT approach developed by Cooper 1993b, a DRS is interpreted not as a function into truth values, but as a function into situations. Another difference is that Cooper uses simultaneous abstraction instead of standard lambda abstraction. This allows argument positions to be distinguished according to their name rather than via their order. The relationship between lambda abstraction and simultaneous abstraction at the level of functions thus parallels that between tuples and partial assignments in the model theory.[4]

2 Focussing on parts of the world

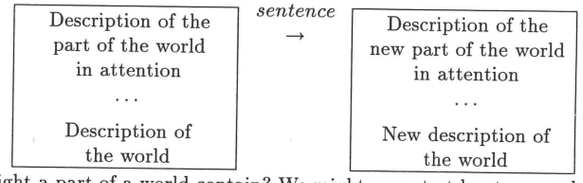
Consider a theory of discourse processing which just takes each sentence and uses it to update a description of the world (or perhaps a description of some imaginary scenario being described) i.e.



It is well known that such a model is inadequate if the description of the world is interpreted as just a set of possible worlds. For example, consider discourses such as the following (due to Barbara Partee):

(9) a. Nine of the ten balls are in the bag. #It is under the sofa.
b. One of the ten balls is not in the bag. It is under the sofa.

The first sentences of the two examples might be expected to rule out the same possible worlds, yet the anaphoric potentials are different. It is data of this kind that led DRT to adopt a more representational level of description. However an alternative is to say that an utterance does not just provide a description of the world, but also brings some part of the world into attention. The first sentence of (9)(a) would bring a part of the world containing 9 balls into attention. The first sentence of (9)(b) would bring a part of the world containing one ball into attention. The model of sentence processing is now as follows:



What might a part of a world contain? We might expect at least some objects and basic relationships. We also won't want to equate a part of a world with a set of possible worlds since we will lose the distinctions we need to make (e.g. we need to distinguish between parts of the world containing 9 balls versus 1 ball).

How does this compare with existing theories? In DPL or PLA, we might describe the contexts as denuded parts of the world: there are entities but no relationships or properties. This description wouldn't be quite correct however, since there is some information kept in the contexts of DPL and PLA which is not contained by parts of the world. In PLA this is the ordering of the entities: we don't have a set of entities but an ordered tuple. Subsequent sections will actually argue

4. Note that STDRT retains the DRT approach of trying to use contexts to provide meanings for the discourse as a whole. Since the meaning of the contexts are not suitable as a meaning of a discourse (they are abstracts), there has to be a step of existential closure occurring at the end of a discourse to turn the abstracts into suitable discourse meanings.

that we need both the relational information supplied by modelling parts of the world, together with ordering (or rather parallelism) information.

How does this compare with DRT? At first sight one would expect there to be little relationship. DRT would seem to provide descriptions of the world rather than parts of the world. However, this is only partly correct. Simple DRS structures can be thought of as modelling parts of the world. In fact Kamp 1981 states that *in fact, formally they are nothing other than partial models, typically with small finite domains.* Simple DRSs also double as a description of the world since the world must be such that the part of the world described by the DRS can be embedded within it.

This dual way of looking at a DRS works for modelling the semantics of indefinite descriptions. However, when we consider generalised quantifiers, the two notions start to diverge. For example, consider the discourses:

(10) Exactly one man came in. He sat down.
(11) A man came in. He sat down.

After processing either of the initial sentences, an appropriate description of the part of the world brought to attention would be a part of the world containing a man who came in. However in DRT we cannot provide the same DRS for both sentences since they differ in their description of the world. Standardly, DRT would build a duplex condition for the first case, with anaphoric binding treated in an E-type fashion. For the second discourse DRT would introduce a new discourse referent for *a man*, with anaphoric binding performed by equating referents. Similarly, for anaphora involving universal quantification, proposals in DRT also tend to be E-type in spirit, with parts of the context copied into the representation for the new sentence (see e.g. Roberts 1987 on telescoping, and Kamp and Reyle 1993 on plurals).

The basis of the treatment proposed in this paper is that contexts consist of the parts of the world (or parts of some imaginary scenario in fictional contexts), which are brought to a hearer's attention. Unlike DRT we allow both indefinites and universal quantifiers to introduce parts of the world into attention. However, to explain the different anaphoric properties of universals and indefinites the contexts are structured differently. Indefinites introduce meta-level disjunctions. For example, if two parts of the world, i.e. two *situations*, **s1** and **s2** both satisfy a sentence containing an indefinite, the context consists of a disjunction of **s1** being in attention, or **s2** being in attention. Universal quantifiers introduce meta-level conjunctions, where more than one situation is brought into attention simultaneously e.g. **s1** is in attention and **s2** is in attention. The proposal resembles DPL as opposed to DRT in working at the equivalent of the model theory rather than at a level of representation.[5] Similar to DPL, information concerning anaphora is not encoded within a representation of sentence meaning, but is provided explicitly by a context.

3 Definites

Most treatments of definite descriptions, whether based on a Russellian or Strawsonian approach, assume some kind of uniqueness constraint[6]. In the Russellian

5. Yet another perspective on the difference between DRT and DPL is to think of the assignments in DPL as pulling out the information which is kept in the structure of the DRS which is relevant for anaphora. Once this information is extracted, we can return to a more traditional semantics. One criticism of this approach is that it doesn't provide the rich structure necessary for a treatment of discourse relations (as in e.g. Asher 1993). However with a situation based account it is perfectly natural to provide rich structuring of the parts of the world in attention instead of putting structure into the description of the world as in DRT.
6. It is worth noting that there are some examples of definites which don't require uniqueness. For example, *Take the lift to the 3rd floor* is fine even if there is more than one lift.

treatments this is built into the meaning of the sentence containing the definite. In Strawsonian treatments, there must be a unique referent if the sentence containing the definite is to get a meaning. However, it is clear that when we utter a sentence such as *the dog ran away* we do not require there to be just one dog in the universe[7]. We can make this intuition clearer by considering a context where some people are looking at a coat and a scarf in a shop window. It is perfectly felicitous to say:

(12) I like **the coat**. Unfortunately it's a bit more expensive than **the one we saw yesterday**.

However, this discourse contains the description **the coat**, yet any universe which satisfies the discourse requires at least two coats.

Barwise and Perry 1983 treated such cases by evaluating definites not in the universe or world as a whole, but in part of the world i.e. some *situation*. Similar motivations led to proposals that definites pick up only salient individuals, or choose from a salient set of individuals (see Cooper 1993a for discussion of McCawley and Lewis's proposals along these lines). In particular, Westerståhl 1984 proposes that determiners generally should be restricted by context sets[8]. It is important to note that we need to allow different noun phrases to pick up different context sets or different situations. Westerståhl makes this point for quantifiers. For example, in:

(13) The English love to write letters. **Most children** have **several pen pals in many countries**

the noun phrase *most children* is presumably restricted to *English children*, an inappropriate restriction for the second noun phrase. Similarly, in (12) we need to use a small context set/situation to evaluate the first definite, and a wider context set/larger situation to evaluate the second.

The second thing to note is that the choice of a situation or context set can be affected by local information within a sentence, not just by global discourse structure. For example, consider the following:

(14) In Forbach, **the castle** is next to a lake. In Blieskastel, **the castle** is on a hill top.

Here *the castle* is evaluated in the two different scenarios brought into salience by the prepositional phrases *In Forbach*, and *In Blieskastel*.

For most semantic theories, it is a far less radical step to incorporate a notion of context set, than a notion of situation, so what are the advantages of a situation based approach? We will concentrate on two phenomena here: bridging reference and state change anaphora. Further arguments for a situation based account of definites are given in Barwise and Perry 1983 and Cooper 1993a.

3.1 Bridging Reference

Anaphoric reference can occur when there is no 'overt' antecedent, but instead the referent is related to a previously mentioned individual, or situation. Such uses are called *bridging reference* (Clark and Haviland 1974). Bridging is sometimes assumed to be limited to cases where an object can be linked to an antecedent noun phrase, or antecedent verb e.g. examples (15) and (16) respectively:

(15) Patience walked into [a room]$_i$. [**The chandeliers**]$_{(i)}$ burned brightly

(16) Keith drove to London. **The car** kept overheating (Garrod and Sanford 1994)

However, it is also common to get cases where it is a particular combination of objects and relationships, plus world knowledge, that allows bridging to occur. For

7. This example is taken from Cooper 1993a.
8. Cooper 1993a similarly argues that determiners generally should be evaluated with respect to situations.

example:

(17) Every Macintosh with a printer has problems with **the connection**

(18) I got home late last night since **the bus** was 20 minutes late

(19) While Mary was washing her hair/clothes, **the shampoo/powder** spilt all over the floor.

(20) No mail came for me today. My secretary was really surprised by **the empty pigeon-hole**

In (17), an example of a donkey sentence with a bridging definite, there is certainly no one antecedent, though we might claim it is some plural object composed out of the computer and its printer. In (18) we use world knowledge that a bus is likely to be involved in someone getting home. (19) was used in experiments by Sanford and Garrod to show that the choice of object for the verb has immediate affect on the definites which are felicitous in the second sentence. This suggests that the anaphora cannot be purely determined by lexical properties of the verb *wash*. In (20) we have an example where the first event does not denote an event, but rather a lack of one.[9]

These examples suggest that the kind of information required in a context includes relations and properties (or events or situations from which this information can be accessed), not just individuals. However, there is a possible way out for context set approaches. This is to claim that the initial processing of the antecedent sentence sets up partially described objects which the bridging definite accesses (see Bos et al. 1995 for a proposal along these lines in DRT). For (19) this seems possible: whenever a washing event occurs one might set up a partially described object corresponding to the washing substance. One could then claim this is what *the shampoo* is anaphoric to.

In general, however, it seems unlikely that such a prediction mechanism can explain the full range of bridging occurrences. Although some cases of bridging may seem primed by an antecedent sentence, in other cases the work seems to be done by the definite (or a combination of both[10]). For example, in (15) it is not particularly likely for a room to contain a chandelier, but highly likely that a chandelier is in a room. Similarly, we can contrast the following examples[11]

(23) a. The police stopped the car since they thought **the trailer/horsebox /roofrack** looked unstable.

b. The police stopped the car since **the dog** looked dangerous.

(a) seems acceptable, even though a car pulling a trailer is relatively uncommon. (b) seems unacceptable, though a car containing a dog is more common. The difference is that trailers are highly likely to be following cars, whereas dogs can be found in far more situations.

Finally it is worth noting that bridging can occur with pronouns as well as definites, providing some motivation for a similar treatment e.g.

(24) oh I was on the bus and **he** didn't stop at the right stop (Brown and Yule 1983)

9. Situation based approaches should generally have no problem with this example, since situations are usually taken to be include the notion of event and state. However, the account we provide later would need extending with a more sophisticated treatment of negation.
10. This point has been made before Groefsema 1995 for slightly different data such as:

(21) Peter: Would you drive a Mercedes?

(22) Mary: I wouldn't drive ANY expensive car (Sperber and Wilson 1986)

Groefsema argues that the assumption that a Mercedes is an expensive car does not become accessible just from Mercedes, but also via expensive car.
11. These examples were inspired by a discussion in Carter 1987. Carter argues against probabilistic methods on the basis of a similar pair. However, he only considers probabilities in the forward direction i.e. how likely it is for a car to be followed by a trailer, not how likely it is for a trailer to be following a car.

(25) My house was broken into last week. **They** took the TV and the stereo (Rich and Knight 1991)

3.2 State-Change Anaphora

Cases involving state-change anaphora are particularly common in task oriented dialogues. For example in recipes we can get anaphora such as:

(26) Mix the flour, yeast and milk. After 5 mins roll **the dough** into thin strips

(27) Melt the gelatine in hot water and fruit juice. Pour **the jelly** into dishes to cool.

Here *the dough* refers to the object which results from mixing the ingredients and then waiting 5 minutes. In a situation based account we could use the described event and the information provided by the definite to infer an appropriate situation which describes the state resulting from the event. For DRT there have already been suggestions to include information about states resulting from events and this might provide the necessary discourse referents (Kamp and Roßdeutscher 1994). However this does suffer from the problems outlined above concerning approaches which rely on prediction. It many cases it is not obvious what the resulting state is, and it is the definite which allows us to infer it. Consider for example:

(28) A tourist wandered into the desert without a compass. I had the unpleasant job of scaring off **the vultures**.

4 Towards a dynamic semantics with situations

In the previous sections we provided evidence to suggest that we need to be able to access relational information from contexts as well as information about entities. In this section we will informally outline a very simple approach to incorporating situations in a dynamic semantics. We will then outline the various problems with this approach which led to the proposal which is given in Section 8.

Reconsider the example sentence *a man buys a picture* in a model where John buys picture 1 and Fred, picture 2. Using tuples as in the PLA system (Dekker 1994) this gives the following context:

{ ⟨ john, picture1 ⟩, ⟨ fred, picture2 ⟩ }

This context encodes the information about the possible sets of (model-theoretic) entities brought to the hearer's attention, and orders the entities.

As argued in the previous section, contexts also need to supply relational information. We can supply this by building contexts such as the following:

{ {buys(john, picture1), man(john), picture(picture1) }
{buys(fred, picture2), man(fred), picture(picture2) } }

Here we provide the equivalent to the model theoretic entities supplied by PLA, but also provide the relations and properties they are involved with. In some versions of situation semantics terminology, *man(john)* would be called an *infon* sometimes written as ⟨man,john,1⟩, and in some versions, a set of infons would constitute a situation, so we would have a context consisting of a set of two situations. The intuitive notion we want to preserve is that the context encodes those parts of the world (or some described scenario) which the conversation is about. We will talk about the context as a set of situations which are parts of the 'actual' situation (or some imagined situation which the conversation is about). The precise status of the contexts, and whether they should really be called situations doesn't seem to be crucial for the discussion of anaphora, so we will remain relatively uncommitted. We will even remain uncommitted as to whether situations should be thought of as

parts of the real world, or merely as hearers' conceptualisations of the world (i.e. mental constructs similar to mental spaces, Fauconnier 1985).

We have now suggested one way in which contexts might look. The second step in providing an account of anaphora is to say how pronouns and definites access the context to obtain an interpretation. Consider something like DPL (Groenendijk and Stokhof 1991). Here the usual assumption for pronouns is that there is some prior coindexing stage in the syntax that tells the pronoun which index (or discourse variable) to pick. For PLA there is no prior coindexing stage, rather we choose dynamically to link a pronoun with the i-th entity from each tuple i.e. the i-th column in the context.

When we consider the contexts built from the kind of situation suggested above, there is nothing equivalent to a notion of column (or corresponding notions of discourse referent or argument position). The only plausible way to access an individual is first to look along a row (i.e. look in a situation), and then choose an individual from that row according to some descriptive content, e.g. by treating *he* as *the male*, *she* as *the female*. This kind of indirect approach seems natural for cases of bridging definites, but is highly controversial for standard anaphoric pronouns (for example, Groenendijk et al. 1995 argue for an indirect approach for definites, but do not commit themselves for pronouns). To see why this issue is so controversial we need to consider the notorious 'two bishops example'[12].

Consider the following sentence in a model where the only bishop meeting is between the two bishops, Cedric and Marcus.

(29) If a bishop meets another bishop he blesses him

A theory using partial assignments such as EDPL (Dekker 1993) creates the following context after processing *a bishop meets another bishop*

{ (x,cedric), (y,marcus) } { (x,marcus), (y,cedric) }

Here we get two different cases, corresponding to replacing the first bishop by *cedric* or by *marcus*. Similarly, in PLA we would get two cases, according to whether we order *cedric* before *marcus* or vice versa.

When we further come to resolve the anaphoric reference, there are again the two cases to consider. Whether we link the *he* and the *him* to the columns labelled by x and y respectively, or vice versa, we will require both Cedric to bless Marcus and Marcus to bless Cedric.

Now consider doing indirect binding, but still with something like EDPL or PLA. In this case, we separately resolve the anaphors in the two cases. This means that we can resolve the *he* to Cedric in the first case, and also to Cedric in the second. The result is that the sentence gets an unintuitive reading which only requires Cedric to bless Marcus, but not for Marcus to bless Cedric.

In the case of a situation based account, things are equally problematic. Presumably any situation which supports Cedric meeting Marcus will also support Marcus meeting Cedric. We therefore might expect to provide an identity relationship between the infons *meets(cedric,marcus)* and *meets(marcus,cedric)*. But this means that the context consists of just the single situation:

{ bishop(cedric), bishop(marcus), meets(cedric,marcus) }

Assuming a treatment of conditionals which checks that all situations supported by the antecedent can be extended to the consequent, this again allows sentence (29) to be true when only Cedric blesses Marcus.

Before rejecting indirect approaches (and, for that matter, situation based approaches), let us consider some linguistic evidence which suggests some advantage in having an indirect approach. We will then try to merge the best attributes of both approaches.

---

12. These kinds of example are due to Jan van Eijck.

5 Advantages of an indirect approach to anaphora

Groenendijk et al. 1995 argue for an indirect approach to dealing with definite anaphora on the basis of examples such as the following:

(30) A man is walking in the park. Another man is walking in the park too. **The tallest man** whistles.

In this example we pick the tallest man from whichever of the men was the tallest. We don't seem to always have to pick the first man, or always the second. This particular example has some problems however. Firstly we could always pick the first, since the two participants are indistinguishable. A slightly different version of the example prevents this:

(31) John bought a pony, a horse and a donkey. He uses **the strongest animal** for ploughing.

Secondly, both of these examples could be criticised on the grounds that superlatives are a special case of definite noun phrases which contain an implicit argument (i.e. we have *the tallest of the men*, and *the strongest of the animals*). To counter these criticisms consider the following example[13]:

(32) When a Sicilian and a Corsican get engaged, **the girl**'s parents are usually more concerned than **the boy**'s.

It is also worth noting that this example works with pronouns i.e.

(33) When a Sicilian and a Corsican get engaged, **her** parents are usually more concerned than **his**.

The cases of conjunction outlined above are problematic for the direct (i.e. column, slot or referent based) approaches to definite anaphora. However, only the last example concerns pronominal anaphora. There are cases, especially cases of disjunction, which seem problematic both for pronominal and definite anaphora using a direct approach. Consider the following examples:

(34) You will be taught by a lecturer or a researcher. **She** will use an overhead projector.

(35) They either took John's car or rented a van. They then drove **it** all around the Highlands.

(36) We can get a young girl to recite a poem, or an older one to read a passage in French. If **it** takes **her** about 10 minutes then ...

(37) Buy a spanner or borrow an adjustable wrench from a local hardware store, but don't use **it** for the small bolts.

Kamp and Reyle (1993:p206) suggest that cases such as (34) can be dealt with by considering *a lecturer* and *a researcher* as alternative characterisations of the same person. They provide a resulting DRS, but leave the construction rules as a topic for further investigation. This kind of approach is however problematic for the other cases. In (35) the potential antecedents do not even share the same role in the first sentence (the car is taken, the van is rented). Instead, what is common between the potential antecedents is that they are both suitable for being driven. In (36) we need to preserve the correct dependencies within each disjunct (which can't be done by introducing two new discourse referents, one of which can be a young girl or an older girl, the other a song or a recitation).

Dynamics semantics also has troubles with some of these examples. Groenendijk and Stokhof 1991 suggest using program disjunction to treat examples similar to (34). Program disjunction unions together the possible output states for each disjunct.The effect is to treat each case separately, with (34) correctly treated as having a semantics equivalent to:

(38) You will be taught by [a lecturer]$_i$ and she$_i$ will use an overhead projector, or you will be taught by [a researcher]$_i$ and she$_i$ will use an overhead projector.

Program disjunction gives the right results for (35) and (36). (37) is more interesting since there are different numbers of indefinites in the two disjuncts. It is not entirely obvious how to achieve the correct results in either total or partial versions of dynamic semantics. We not only have to merge the different cases into a single context, where appropriate individuals appear in appropriate columns, but also must ensure that discourse referents which were created by only one disjunct are inaccessible outside that disjunct. PLA could provide a neat treatment of examples (34), (35) and (36) since there is no need to provide prior coindexing. However there would be particular problems with (37) since, even if the tuples could be merged, the *it* would have to refer to the spanner and to the store rather than to the spanner and to the wrench, since the spanner and the store are equidistant from the pronoun.

Treating these kinds of disjunction[14] within the simple situation based model outlined above is particularly simple. Similar to Groenendijk and Stokhof, the different outputs can be unioned together. However, each output is just a situation, and there is no problem with unioning together situations of different sizes. The use of an indirect approach to anaphora also means that there is no need to allocate each entity in separate disjunctions into some 'correct' column, since each situation is considered independently. We will consider disjunction more formally in Section 8.

6 Between a direct and indirect approach to anaphora

The discussion so far seems to suggest that a direct approach to anaphora is necessary for the 'bishop' example, but that indirect approaches might have advantages for cases of disjunction, cases involving *the tallest* etc., and, from our earlier discussions, cases of bridging reference where the antecedent can't easily be predicted. How can we reconcile the data?

One feature of the 'bishop' example is that both pronouns are ambiguous. In contrast, in the other examples considered, the pronouns or definites were unambiguous: there was just a single individual satisfying the description in each case. It thus seems that either an anaphor picks up a unique individual in each individual case, or it picks parallel individuals.

How can we build in a notion of parallelism into a situation based account? One relatively simple way is to associate individuals with indices. If, for example, two individuals in different situations were both associated with the same indefinite noun phrase, then they share the same index. The situations now not only encode the part of the world brought into attention, but also something of the way in which each individual in the situation was introduced into attention. For example, interpretation of the antecedent of the conditional in the 'bishops' example would set up two situations with different indexing i.e.

{ bishop(cedric), index(cedric,1), bishop(marcus), index(marcus,2), meets(cedric,marcus) }

{ bishop(marcus), index(marcus,1), bishop(cedric), index(cedric,2), meets(cedric,marcus) }

13. This example is again not entirely watertight since one could argue that this is a case of bridging to the verb *engaged*. However, bridging itself is certainly a problem for simple column based accounts since there is no column for the anaphor to link to.

14. This is just one class of disjunction. Other kinds of example include *I expect a man or a woman will come. She would be happy to help.* In this example we only need to consider extending the situations introduced by a woman coming. The second sentence says nothing more about the situations introduced by a man coming.

As before we look into each situation separately, but now an object can be chosen via its index.

To obtain the two ways in which a pronoun can refer, we interpret a pronoun either as either a simple definite, or as a definite with a restriction that the individual must have the appropriate index. Using generalised quantifier notation, the two interpretations for *he sleeps* would be:

- the(λx.male(x), λy.sleeps(y))
- the(λx. male(x) & x = $\widehat{i}$, λy.sleeps(y))

$\widehat{i}$ is interpreted in a situation as an object with index i. In the formalism we present in Section 8, the second of these formulae is effectively equivalent[15] to the simpler:

male($\widehat{i}$) $\wedge$ sleeps($\widehat{i}$)

This is due to the fact that indices are chosen so that no two items in the same situation share the same index[16].

We have suggested that pronouns can be interpreted by simple definites (in contrast to the elaborate definites provided by E-type approaches). How should definites be interpreted? Here we will give two pieces of data which have affected our approach, without giving any justice to the huge literature. The first piece of data is examples such as:

(39) John saw a doctor at the health centre. **She/the woman** was very helpful.

Here we seem to get a correct interpretation for the discourse if we simply rule out the cases from the first sentence which involve male doctors. This suggests that there is no presupposition failure involved, just a sentence coming out false in certain cases.

The second piece of data is from definites involving other noun phrases e.g.

(40) The man who owned a donkey brought **it** to the fair

Here the *it* can access its antecedent, *a donkey*, even though this is nested within the definite describing the man. This is problematic for treatments such as Barwise and Perry's, where the descriptive content of definites is thrown away once the referent is found. Consider the following sentence:

(41) The man in the red vest is the governor

Barwise and Perry convincingly argue that the the descriptive content of the definite should not be added to the content of the sentence as a whole. The sentence does not seem to commit the speaker to having asserted that there is a man in a red vest. For example, it would be infelicitous for the hearer to answer: *It's not red, it's pink.* Our treatment similarly does not add the content of a definite to the content of the sentence. However, we do add it to the context i.e. the speaker brings the information into attention as opposed to asserting it. This allows treatment of (40).

A further question concerns whether definites themselves ought to get a reading similar to the indexing reading of the pronoun. Even with the limiting of uniqueness to uniqueness within a resource situation, we have too strong a requirement for some anaphoric definites. Consider the following sentence:

(42) When a Protestant marries a Catholic, **the vicar** is careful about the words he chooses

Consider if in one case the Protestant just happens to also be a vicar. The sentence seems to apply equally well to this case, with no possibility of interpreting *the vicar* as *the Protestant.* However, in indirect approaches such as Groenendijk, Stokhof and Veltman's, and in what has been proposed above, this is exactly what might happen, causing uniqueness to fail. Moreover, this is an example where there is no

15. There is a minor difference in that the definite introduces an extra redundant index.
16. We might want to change this. For example, coordinated noun phrases perhaps should result in the individuals concerned sharing the same index.

syntactic coindexing of *the vicar* with an antecedent, so the lack of a reading couldn't be enforced syntactically (as, for example, other cases of parallelism are enforced by assuming prior coindexing in DPL)[17]. The view of definites as also being ambiguous between two readings, one involving indexation one not, might gain support from considering plural anaphora. Firstly, there are convincing arguments (e.g. Elworthy 1995) that plural pronouns require the kind of referential treatment we have proposed here (i.e. a treatment which considers the actual sets of entities that pronouns refer to, as opposed to some claimed syntactic antecedent). Secondly, there are cases of plural anaphora which seem to exhibit the same bifurcation between consideration of all individuals in a context versus considering only individuals with the same index. Contrast the following examples:

(43) John met Peter and Susan. **The boys** then went off to the cinema.
(44) Everyone at the school went to visit the other children. **The boys** really enjoyed the visit

In the first, example, *the boys* refers to all the boys in the context without worrying about their indexing. In contrast, in the second example, *the boys* has additional readings where it refers to all the boys with the same index (i.e. all the boys who visited, or all the boys who were visited).

7 Some final motivation: plurals and telescoping

In the last section we suggested that a quantifier such as *everyone* introduces individuals which share the same index. But what kind of context should these individuals be embedded in? The anaphoric properties of universal quantifiers are very different from those of indefinites, so we wouldn't want to produce the same kinds of context for both (i.e. sets of situations). However, if we don't use sets, what are the other options?

Let us return to the original intuitions we used for introducing sets of situations. The idea was that each situation corresponds to a part of the world brought to attention. In the case of indefinites, we may have more than one part of the world to consider: the hearer knows that one of them is in attention, but doesn't know which one. We thus have a kind of meta-level disjunction e.g. (s1 is in attention) or (s2 is in attention). Given this perspective, the obvious kind of context to create for a universal quantifier is one where various situations are brought into attention in parallel e.g. (s1 is in attention) and (s2 is in attention) i.e. a meta-level conjunction.

Consider how this might help with a treatment of plural anaphora to universal quantifiers. Given the discourse,

(45) Every child had a good time visiting the grotto. **They** particularly enjoyed the reindeer

the first clause creates a set of situations which are all in attention. The referent for the pronoun is then created by summing parallel individuals (i.e. individuals sharing the same index) in each situation. Elworthy 1995 works out a similar approach in far more detail. In Elworthy's work, the first clause would set up a set of tuples of individuals, and the *they* would sum together all individuals sharing a particular column. Note that this approach differs from Dekker's PLA system, although the contexts look similar. Elworthy uses sets of tuples as the result of universal quantification, in cases where we would use conjunctive situations. Dekker uses sets of

17. Extending coindexing to bridging suggests that if two parallel situations can be extended by the same bridging inference, then the resulting situations include parallel individuals with the same index. In this respect the treatment resembles DRT, where an inference can set up a new discourse referent. There is no stipulation that discourse referents have to be linked to noun phrases introduced in the discourse itself.

tuples as the result of existential quantification, where we would use disjunctive situations.

The idea of keeping multiple situations in parallel attention also provides a neat treatment of telescoping phenomena. Consider the following examples:

(46) Each degree candidate walked to the stage. **He** took **his** diploma from the Dean and returned to **his** seat[18]

(47) Every girl is expected to swim three lengths. **She** is also expected to stay under water for 5 minutes.

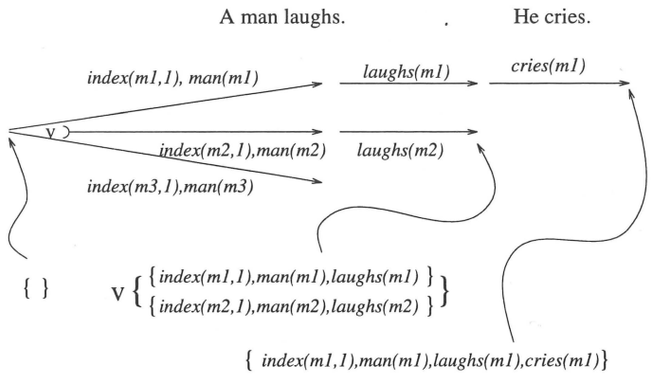
(48) Every chess set comes with a spare pawn. **It** is taped to the top of the box (Sells 1985)

In each of these we can think of the first sentence as setting up a set of situations which are simultaneously in attention. The effect of the second sentence is just to extend each of the situations. We will give formal rules to achieve this in the next section.

8 A Dynamic Semantics with Situations

The discussion above could lead to a variety of different formalisations. The formal system we shall describe here should be seen as just one example. The prime consideration here has been simplicity, and we will suggest some refinements in the conclusion. We assume, similar to DPL, that we are starting with a language of generalised quantifiers rather than directly with natural language. This assumption is justified by the fact that there are already well known techniques for converting reasonable sized fragments of natural language to a language of language of GQs e.g. using 'Montague-style' techniques. A direct mapping from natural language would be more desirable, but introduces extra complexity.

The rules for the simple fragment are given in Figure 1. Let us first run through some examples before looking at specific features of some of the rules. Consider the discourse *A man laughs. He cries.* Now consider an actual situation in which there are three men, **m1**, **m2** and **m3**. **m1** and **m2** laugh, but only **m1** both laughs and cries. The result of updating an initially empty situation is given below:



The rule for existential quantifiers picks a new index, index 1, and forms a disjunction from the situations resulting from extending the situation, first with the indexed object, then with $\mathbf{man}(\hat{1})$ followed by $\mathbf{laughs}(\hat{1})$. $\hat{i}$ is a term in the logical language which is interpreted in a situation as the object with the index **i**. A disjunctive situation can be extended by extending each member of the disjunction.

18. This example is given in Roberts 1987 acknowledging Barbara Partee.

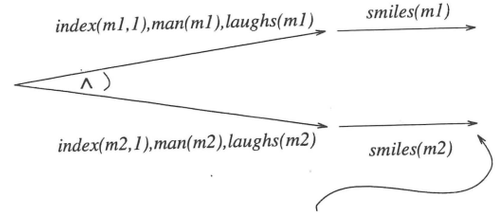
The result is another disjunctive situation, possibly with some cases pruned out (or extra cases introduced by new indefinites). This exactly parallels what occurs in e.g. EDPL with sets of possible assignments.

Now consider the sentence *If a man laughs, he smiles* in an actual situation similar to the one above, but with the added feature that **m1** and **m2** both smile. The implication rule is based on the relatively standard idea that all possible situations satisfying the antecedent must be able to be extended by the consequent. The exact formulation of the rule is however unusual. In the case where the antecedent creates a disjunctive situation, the effect is to turn the disjunction into a conjunction which is then extended by the consequent. These two steps are shown in the following figure:

If a man laughs

$$\wedge \{ \{ index(m1,1),man(m1),laughs(m1) \} \{ index(m2,1),man(m2),laughs(m2) \} \}$$

If a man laughs he smiles



$$\wedge \{ \{ index(m1,1),man(m1),laughs(m1),smiles(m1) \} \{ index(m2,1),man(m2),laughs(m2),smiles(m2) \} \}$$

Since we have a conjunction, both of the situations satisfying the antecedent must be extended (otherwise we would get a $\perp$ into the conjunction, which is equivalent to the whole mapping to $\perp$).

The motivation for this particular formulation of the implication rule comes from looking at *suppose*. Consider the following discourses:

(49) Suppose Jones owns a book on semantics. Then he uses it (Kamp and Reyle 1993)

(50) Suppose Jones has a daughter. Then will she help him?

Consider the first[19]. Here we seem to need to consider every scenario in which Jones owns a book on semantics. Interpreting *John owns a book on semantics* will create a (meta-level) disjunction of situations e.g.

$$\vee \{ \{ owns(john,book1),\ book\text{-}on\text{-}semantics(book1) \} \{ owns(john,book2),\ book\text{-}on\text{-}semantics(book2) \} \}$$

The effect of *suppose* is to turn this from a disjunction into a conjunction i.e. it brings all the cases into parallel attention. We interpret *then he uses it* in this conjunctive context, giving the interpretation that he uses all the books on semantics which he owns.

19. I would prefer this example if *uses* were replaced by would use. Modals often seem to signal that we are remaining in a hypothetical context. Once we use a non-modal this signals that we should look again at the cases rejected by the hypothesis as well as the case considered under the hypothesis. We will give no account here of how to close off a hypothetical context.

This treatment of implication gives strong readings to donkey sentences. For example, the sentence:

(51) If a farmer owns a donkey he beats it

is interpreted as true (provides a transition to something other than $\bot$) only if every case in which a farmer owns a donkey can be extended to one in which he beats it. There are various ways in which we might attempt to get weak readings. One is to retain more structure so that a disjunction of a disjunction is not flattened to a disjunction of the individual cases. Asymmetric readings would result from only converting the top levels of disjunctions into conjunctions.[20] Another possibility is to give a singular pronoun such as *it* an extra reading corresponding to *one of them.*

Let us now look at some individual features of the rules in Figure 1. Figure 1 uses the convention **s** for situations, **ss** for situations or structured situations, **as** for the actual situation, and **rs** for the resource situation.

$$s[[R\ t_1\ ...\ t_n]]_{as} = s \cup \{R(t_1,...,t_n)\}$$
$$\text{if } [[t_j]]^s_{as} = t_j,\ [[R]]^s_{as} = R, \text{ and } as \models R(t_1,...,t_n)$$
$$= \bot \text{ otherwise}$$

$$s[[\neg\phi]]_{as} = s$$
$$\text{if } s[[\phi]]_{as} = \bot$$
$$= \bot \text{ otherwise}$$

$$s[[\phi \wedge \psi]]_{as} = (s[[\phi]]_{as})[[\psi]]_{as}$$

$$s[[\phi \vee \psi]]_{as} = \bigvee\{ss \mid ss \in_V s[[\phi]]_{as} \vee ss \in_V s[[\psi]]_{as}\}$$

$$s[[\phi \rightarrow \psi]]_{as} = (\bigwedge\{ss \mid ss \in_V s[[\phi]]_{as}\})\ [[\psi]]_{as}$$

$$s[[\exists_i(\phi,\psi)]]_{as} = (\bigvee\{ss \mid \exists obj \in as.\ ss \in_V (s \cup \{index(obj,i)\})[[\phi(\hat{i})]]_{as}\})\ [[\psi(\hat{i})]]_{as}$$

$$s[[\forall_i(\phi,\psi)]]_{as} = (\bigwedge\{ss \mid \exists obj \in as.\ ss \in_V (s \cup \{index(obj,i)\})[[\phi(\hat{i})]]_{as}\})\ [[\psi(\hat{i})]]_{as}$$

$$s[[pro_{i*}(\phi,\psi)]]_{as} = (s[[\phi(\hat{i})]])[[\psi(\hat{i})]]$$

$$s[[pro_i(\phi,\psi)]]_{as} = s[[the_i(\phi,\psi)]]_{as}$$

$$s[[the_i(\phi,\psi)]]_{as} = (\bigvee\{ss \mid \exists obj \in rs.\ ss \in_V (s \cup \{index(obj,i)\})\ [[\phi(\hat{i})]]_{rs}\ \&$$
$$\forall obj' \in s \text{ s.t. } obj' \neq obj.\ (s \cup \{index(obj,i)\})\ [[\phi(\hat{i})]]_{rs} = \bot$$
$$\})\ [[\psi(\hat{i})]]_{as}$$

$$\bigvee S[[\phi]] = \bigvee\{ss' \mid \exists ss \in S.\ ss[[\phi]]ss'\} \qquad \bigwedge S[[\phi]] = \bigwedge\{ss' \mid \exists ss \in S.\ ss[[\phi]]ss'\}$$

20. This makes asymmetric quantification dependent upon the scoping of the quantifiers in the antecedent. The result is that outerscope quantifiers in the antecedent correspond to the cases which are quantified over. Kadmon 1988 argues that this is the correct generalisation. She gives the remaining quantifiers a weak (existential) reading. However, there is some data which doesn't seem to fit this. Consider *If a poor girl has a toy she usually gets good use from it.* Here there seems to be a reading where most poor girls play with *most* of their toys. Zimmerman 1993 makes a similar point with regard to the sentence *Most environmentalists who buy a car buy a pink diesel.*

$$\bigwedge(A \cup \{\bot\}) = \bot \qquad \bigvee(A \cup \{\bot\}) = \bigvee A$$

$$\bigwedge\{A\} = A \qquad \bigvee\{A\} = A$$

$$\bigvee\{\} = \bot \qquad \bigwedge\{\} = \top$$

$$[[\hat{i}]]^s_{as} = obj \text{ if } s \models index(obj,i)$$

$A \in_V ss$ iff
$\exists S.\ ss = \bigvee S$ and $A \in_V S$ or
$ss = s$ for some s and $A = s$ or
$\exists S.\ ss = \bigwedge S$ and $A = ss$

Figure 1: A Dynamic Situation Semantics

The first rule interprets a ground term in the logical language. If the arguments are all constants, the result is to add the corresponding infon into the situation e.g.[21]

{ *sleeps(john)* } [[sleeps(harry)]] { *sleeps(john), sleeps(harry)* }

The rule for negation is a rule of negation by failure similar to DPL (see e.g. Krahmer 1994 for a more sophisticated treatment of negation which allows anaphoric reference in cases of double negation). Note that here we have already distributed down to an individual situation. If, for example, we had started with $\bigvee\{\mathbf{s1},\mathbf{s2}\}$, and if $\mathbf{s1}[[\phi]]\bot$ and $\mathbf{s2}[[\phi]]\mathbf{s3}$, then we would get $\bigvee\{\mathbf{s1},\mathbf{s2}\}[[\phi]]\mathbf{s1}$ by using the rule for distributing into a disjunctive situation.

The rule for conjunction is again similar to the rule in DPL i.e. first process ϕ, then process ψ.

The rule for disjunction provides a meta-level disjunction over the union of the various situations or structured situations (excluding $\bot$), which satisfy either of the disjuncts. This is similar to unioning sets of assignments as suggested by Groenendijk and Stokhof 1991, but has the advantage that there is no problem using meta-level disjunction for situations containing different numbers of indefinites. The similar treatment of disjunctions and indefinites means that the following sentences get similar semantics if evaluated in an actual situation containing just two friends of *Mary*, *John* and *Bill*:

(52) a. Everyone who met a friend of Mary's liked him
b. Everyone who met John or Bill liked him

The rule for implication was discussed earlier. Note that in the case where the antecedent of a conditional is not supported by the actual situation, we get a conjunction over an empty set. Evaluating the consequent in a context consisting of a conjunction over an empty set just returns the same context (according to the rule for extending a conjunctive situation). We call a conjunction of an empty set $\top$. For any ϕ, $\top[[\phi]]\top$ (similarly, for any ϕ, $\bot[[\phi]]\bot$). A further sentence will also successfully extend this context. This is not necessarily correct. Sometimes we will want to continue to extend the possible contexts set up by the assumption of an implication, but not always. There need to be other rules to allow a conjunctive situation to be closed off, perhaps to form a single situation consisting of the set of all infons in any of the conjoined situations.[22]

The rule for existential quantification was discussed earlier. The rule should be refined to allow existential quantifiers to take wide scope in telescoping cases (consider e.g. (46) with *the Dean* replaced by *a distinguished visitor*). The rule for universal quantification is similar to that for implication. This is made possible by

21. In the examples we use the convention that **john**, a constant in the language of GQs, is interpreted as the individual *john* in the actual situation. The actual situation takes the role that a model would in a standard model theoretic interpretation.
22. There are issues here concerning the context prior to the implication. For example, we might want to let the cases which didn't satisfy the supposition be continued. Further consideration of negation, and perhaps modality might be useful here.

the use of the $\widehat{\ }$ convention as opposed to the standard approach in e.g. First Order Logic of using variables plus an assignment function. As with implication, a false restrictor gives the context $\top$ which can be successfully extended by any new input. This is correct for telescoping cases, but as mentioned earlier, not correct in all cases. The rule should also be refined to allow for restricted quantification i.e. choice of objects not from the actual situation as a whole but from some resource situation.

Finally there are the rules for pronouns and definites. We assume that pronouns in the input language are in generalised quantifier format e.g. *he sleeps* is represented as:

$\text{pro}_{i/i*}(\lambda x.\text{male}(x),\lambda y.\text{sleeps}(y))$

There are two interpretation rules as discussed earlier. The first plugs in the $\widehat{i}$ term into a conjunction of the restrictor and scope. The second rule for pronouns treats them as definites.

For definites, we want to fetch the unique item in the resource situation which satisfies the restrictor, and to thread the context formed by evaluating the restrictor into the scope (to deal with the embedded noun phrases mentioned earlier). This is achieved by threading the context formed by the one object in the situation which doesn't map to a null context. Currently there are two choices for the resource situation: the situation in which the definite is being evaluated, or the actual situation. If the contexts were further structured using discourse relations this would introduce other possibilities.

Finally there are rules for the $\bigvee$ and $\bigwedge$. $\bigvee$ and $\bigwedge$ are defined as operations over sets, but we could equally have used a binary *meet* and *join* operation with properties of idempotence, commutativity and associativity. The resulting operations would then resemble standard lattice operations. It is however questionable whether we would want to make contexts into lattices since this would destroy some necessary structure. For example, the consequence of adding a partial ordering appropriate to a lattice structure would be to require that $\bigvee\{s1, \bigwedge\{s1, s2\}\} = s1$. However it is plausible that a plural pronoun might distinguish between these structures.

9 Conclusions and Future Directions

This paper has provided a preliminary account of how we can construct contexts consisting of conjunctively and disjunctively structured situations (including the possibility of disjunctions of conjunctions of disjunctions etc.), and how this can help in dealing with various kinds of anaphora.

So far, the structuring of the context has been according to the logical properties of the utterances. What we have not done is to structure the discourse in the sense that is normally meant by discourse structure. This requires structuring in the orthogonal dimension i.e. each individual situation needs to be structured according to the way the discourse is segmented and how the discourse relates to what it describes (see e.g. Asher 1993 and Poesio 1993 for approaches to incorporating discourse structuring into DRT and a version of situation semantics respectively). Once such structuring were added, it would make sense to allow a greater freedom in the choice of resource situation. For example, we might allow salient sub-situations of the current discourse situation in addition to the discourse situation itself and the actual situation.

Will this greater freedom of choice allow us to cope with all cases? The first kind of case which might be problematic is represented by example (14) repeated here:

(53) In Forbach, **the castle** is next to a lake. In Blieskastel, **the castle** is on a hill top.

How are the resource situations chosen? In this example, there seems no need for the 'Forbach' situation (the major buildings and events associated with the town Forbach) to be salient prior to the sentence being uttered. Instead, we seem to be able to carve it out of the actual situation using the content of the utterance itself.

There are also some cases involving pairs of definites which suggest that you cannot evaluate each definite separately in a resource situation. Consider the following sentences[23], uttered in a context where there are two boys, two dogs, but only one boy holding a dog.

(54) **The boy** is holding **the dog** *too* firmly.

(55) **The boy** is really enjoying holding **the dog**.

If we try to evaluate one of the definites by itself we will get uniqueness failure. However, the sentences seem perfectly felicitous in this context. It seems that we need to pick out the pair of a boy and a dog such that the boy is holding the dog, and then assert that he is holding it too firmly, or that he is really enjoying holding it i.e. use a sentence's given, or, perhaps, its presuppositional information together with a resource situation to establish the referents. This suggests that we need to extend the approach taken in this paper to allow even more interaction between discourse context and sentential context.

References

Asher, N.: 1993, *Reference to Abstract Objects in Discourse*, Kluwer, Dordrecht, Dordrecht

Barwise, J.: 1987, Noun phrases, generalized quantifiers and anaphora, in P. Gärdenfors (ed.), *Generalized Quantifiers: linguistic and logical approaches*, pp 1–30, Reidel, Dordrecht

Barwise, J. and Perry, J.: 1983, *Situations and Attitudes*, MIT Press, Cambridge Mass, Cambridge, MA

Bos, J., Buitelaar, P., and Mineur, A.-M.: 1995, Bridging as coercive accommodation, in *Proceedings of Computational Logic for Natural Language Processing (sponsored by Elsnet, Compulog-net, Eagles)*, South Queensferry, Scotland

Brown, G. and Yule, G.: 1983, *Discourse Analysis*, CUP, Cambridge

Carter, D.: 1987, *Interpreting Anaphors in Natural Language Texts*, Ellis Horwood, Chichester, Sussex

Clark, H. and Haviland, S.: 1974, Psychological processes as linguistic explanation, in D. Cohen (ed.), *Explaining Linguistic Phenomena*, Hemisphere Publication Corporation, Washington

Cooper, R.: 1993a, Generalised quantifiers and resource situations, in P. Aczel, D. Israel, Y. Katagiri, and S. Peters (eds.), *Situation Theory and its Applications, v.3*, Chapt. 8, pp 191–212, CSLI and University of Chicago, Stanford

Cooper, R.: 1993b, Towards a general semantic framework, in R. Cooper (ed.), *DYANA-2 Deliverable R2.1.A*, pp 49–97, The DYANA-2 Project Administrator, ILLC, University of Amsterdam

Dekker, P.: 1993, *Transsentential Meditations*, Ph.D. thesis, Department of Philosophy, University of Amsterdam

Dekker, P.: 1994, *Predicate Logic With Anaphora*, Technical Report LP-94-11, ILLC, Amsterdam

Does, J. v. d.: 1993, The dynamics of sophisticated laziness, in J. Groenendijk (ed.), *DYANA-2 Deliverable R2.2.A, Part I*, The DYANA-2 Project Administrator, ILLC, University of Amsterdam

23. These sentences have some similarity with the cases of embedded definites examined in Haddock 1988 such as *the rabbit in the hat sleeps*. This seems acceptable in a context where there are two hats, provided only one has a rabbit in it.

Does, J. v. d.: 1994, Formalizing E-type anaphora, in P. Dekker and M. Stokhof (eds.), *Proceedings of the Ninth Amsterdam Colloquium*, pp 229–248, ILLC, Amsterdam
Elworthy, D.: 1995, A theory of anaphoric information, *Linguistics and Philosophy* 18(3), 297–332
Evans, G.: 1980, Pronouns, *Linguistic Inquiry* 11, 337–362
Fauconnier, G.: 1985, *Mental Spaces*, MIT Press, Cambridge, Mass.
Fernando, T.: 1994, What is a DRS?, in R. Cooper and J. Groenendijk (eds.), *DYANA-2 Deliverable R2.1.B*, The DYANA-2 Project Administrator, ILLC, University of Amsterdam
Garrod, S. and Sanford, A.: 1994, Resolving sentences in a discourse context, in M. Gernsbacher (ed.), *Handbook of Psycholinguistics*, Academic Press
Groefsema, M.: 1995, Understood arguments: A semantic/pragmatic approach, *Lingua* 96, 139–161
Groenendijk, J. and Stokhof, M.: 1991, Dynamic predicate logic, *Linguistics and Philosophy* 14, 39–100
Groenendijk, J., Stokhof, M., and Veltman, F.: 1995, Coreference and contextually restricted quantification, in J. Groenendijk (ed.), *DYANA-2 Deliverable R2.2.C*, The DYANA-2 Project Administrator, ILLC, University of Amsterdam
Haddock, N.: 1988, *Incremental Semantic Interpretation and Incremental Syntactic Analysis*, Ph.D. thesis, University of Edinburgh
Heim, I.: 1982, *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. thesis, University of Massachusetts, Amherst
Heim, I.: 1990, E-type pronouns and donkey anaphora, *Linguistics and Philosophy* 13(2), 137–177
Kadmon, N.: 1988, Asymmetric quantification, in J. Groenendijk, M. Stokhof, and F. Veltman (eds.), *Proceedings of the Sixth Amsterdam Colloquium*, ILLC, University of Amsterdam
Kadmon, N.: 1990, Uniqueness, *Linguistics and Philosophy* 13(3), 273–342
Kamp, H.: 1981, A theory of truth and semantic representation, in J. G. et al. (ed.), *Formal Methods in the Study of Language*, Mathematisch Centrum, Amsterdam
Kamp, H. and Reyle, U.: 1993, *From Discourse to Logic*, Kluwer, Dordrecht
Kamp, H. and Roßdeutscher, A.: 1994, Remarks on lexical structure and DRS-construction, *Theoretical Linguistics* 20(2/3), 97–164
Krahmer, E.: 1994, Partiality and dynamics; theory and application, in P. Dekker and M. Stokhof (eds.), *Proceedings 9th Amsterdam Colloquium*, pp 391–410, ILLC, Amsterdam
Poesio, M.: 1993, A situation-theoretic formalization of definite description interpretation in plan elaboration dialogues, in P. Aczel, D. Israel, Y. Katagiri, and S. Peters (eds.), *Situation Theory and its Applications, v.3*, pp 339–374, CSLI and University of Chicago, Stanford
Rich, E. and Knight, K.: 1991, *Artificial Intelligence*, McGraw Hill
Roberts, C.: 1987, *Modal Subordination, Anaphora, and Distributivity*, Ph.D. thesis, University of Massaschusetts
Sells, P.: 1985, *Restrictive and Non-Restrictive Modification*, Technical Report CSLI-85-28, Centre for Study of Language and Information, Stanford
Sperber, D. and Wilson, D.: 1986, *Relevance: Communication and Cognition*, Blackwell, Oxford
Westerståhl, D.: 1984, Determiners and context sets, in J. v. Benthem and A. t. Meulen (eds.), *Generalized Quantifiers in Natural Language*, Foris, Dordrecht
Winograd, T.: 1976, *Understanding Natural Language*, Academic Press, New York
Zimmerman, E.: 1993, Paul Dekker on the proportion problem, in R. Cooper (ed.), *DYANA-2 Deliverable R2.1.A*, pp 47–48, The DYANA-2 Project Administrator, ILLC, University of Amsterdam

In situ binding: a modal analysis

Michael Moortgat (OTS)

Abstract
In this paper we compare two multimodal deconstructions of the *in situ* binder $q(A, B, C)$, proposed in [9] for the scoping of generalized quantifiers. The wrapping analysis of [13] is shown to be of limited generality: it restricts the occurrence of generalized quantifier expressions to *associative* environments — environments where sensitivity for constituent structure is lacking. We propose an alternative deconstruction where the wrapping operation is independent of resource management assumptions about the structural context. The analysis is based on the general theory of structural control proposed in [8]: the interaction principles for the wrapping operation are fine-tuned in terms of unary modal control devices.

1 Resource control

In [10, 8], the categorial vocabulary is extended with a pair of unary modal operators $\Diamond, \Box^{\downarrow}$. The base logic for the unary operators is captured by the residuation laws that also govern the familiar binary vocabulary, cf. (1) below.

$$\begin{array}{c} \Diamond A \to B \quad \textit{iff} \quad A \to \Box^{\downarrow} B \\ A \to C/B \quad \textit{iff} \quad A \bullet B \to C \quad \textit{iff} \quad B \to A \backslash C \end{array} \tag{1}$$

The embedding theorems of [8] establish a general theory of communication for substructural grammar logics: in interaction with the standard binary multiplicatives, the unary vocabulary provides full control over structural organization in terms of precedence, dominance and dependency. The control allows both for the imposition of structural constraints in regimes with a flexible resource management, and for licensing structural relaxation in regimes with a more stringent structure sensitivity. This result suggests a natural division of labour between 'syntax' and 'semantics', with **LP** playing the role of the default semantic composition language, and the pure residuation logic **NL** the default language of structural composition. The intermediate territory can be navigated by means of the modal control operators.

We apply this method to a painful open question in the categorial literature: the proper characterization of an *in situ* binding operation such as required for the scopal possibilities of generalized quantifiers. In §2, we discuss the tension between semantic expressivity and syntactic discrimination as it arises in the context of quantifier scoping. In §3, we present the analysis of 'in situ' binding proposed in [13], which exploits the properties of an associative default resource management regime. In §4, we develop a generalized multimodal analysis which is independent of resource management assumptions for the environment the generalized quantifier expressions find themselves in.

2 Quantifier scope: syntax-semantics tension

Consider generalized quantifier expressions like 'someone', 'everybody'. ¿From the perspective of **LP**, we can study their contribution to natural language meaning via a standard Fregean type assignment $(e \to t) \to t$, with lexical recipes $\lambda x.\exists y[x(y)]$, $\lambda x.\forall y[x(y)]$. The **LP** notion of derivability, of course, is too crude to offer a unified deductive account of the composition of meaning *and* form. Suppose we want to refine the **LP** type $(e \to t) \to t$ to take syntactic fine-structure into account. Within **L**, one can find two directional realizations of $(e \to t) \to t$ compatible with the fact that generalized quantifiers occupy the positions of ordinary (proper noun) noun phrases: $s/(np\backslash s)$ and $(s/np)\backslash s$. But imposing order sensitivity in the type-assignment already causes the loss of scope readings one wants to preserve. Compare 'peripheral' versus 'medial' occurrences of generalized quantifiers. Given a 'direct object' assignment $(s/np)\backslash s$ to 'someone', both the (a) and (b) readings of (2) are **L**-derivable. Given a 'subject' assignment $s/(np\backslash s)$ the (c) reading is not derivable: in **L** one only derives the narrow scope reading (d).

$$\begin{array}{lll} (a) & \text{A thinks B loves someone} & \rightsquigarrow \text{someone}(\lambda x.\text{thinks}(\text{loves}(x)(\text{b}))(\text{a})) \\ (b) & & \rightsquigarrow \text{thinks}(\text{someone}(\lambda x.\text{loves}(x)(\text{b})))(\text{a}) \\ (c) & \text{A thinks someone loves B} & \rightsquigarrow \text{someone}(\lambda x.\text{thinks}(\text{loves}(\text{b})(x))(\text{a})) \\ (d) & & \rightsquigarrow \text{thinks}(\text{someone}(\lambda x.\text{loves}(\text{b})(x)))(\text{a}) \end{array} \tag{2}$$

The diagnosis of the problem is not difficult: the (2c) reading would require the generalized quantifier expression to enter into structural composition with a *discontinuous* configuration of resources: such syntactic behaviour is beyond the expressivity of the **(N)L** connectives:

$$\boxed{\text{A thinks}}\ \text{someone}\ \boxed{\text{loves B}}$$

Below, we briefly compare two strategies to resolve this problem: the *rule-based* approach of [4] which postulates type-change axiom schemata to regain the lost readings, and the *deductive* approach of [9] which enriches the vocabulary of connectives with a connective for 'in situ' binding which turns the axiom schemata of [4] into derivable theorems.

FLEXIBLE MONTAGUE GRAMMAR. The proposal of [4] is formulated as a flexible version of Montague Grammar (FMG). Syntactically, FMG is restricted to combine phrases by means of function application rule schemata. In order to accommodate quantificational scope ambiguities, the category-to-type mapping is relaxed to a relation rather than a function: a given syntactic type is associated with a set of semantic types. The semantic types are not unrelated: from a generator type an infinite number of semantic types (and the associated meaning recipes) are derived via the type-shifting rule schemata of Value Raising (VR), Argument Lowering (AL), and Argument Raising (AR).

Let us identify the pure application syntax of FMG as **NL**, and try to pinpoint exactly where the type-shifting schemata give a surplus inferential capacity. Lifting (3a), and its generalizations via (3b,c) to Value Raising and

Argument Lowering are universally valid already in the pure residuation logic **NL**: they reflect the monotonicity properties of the implicational type constructors.

$$\begin{array}{ll} (a) & A \to B/(A\backslash B), \quad A \to (B/A)\backslash B \\ (b) & \text{if } A \to B, \text{ then } A/C \to B/C \text{ and } C\backslash A \to C\backslash B \\ (c) & \text{if } A \to B, \text{ then } C/B \to C/A \text{ and } B\backslash C \to A\backslash C \end{array} \tag{3}$$

Argument Raising, as a semantic type-shifting rule, is schematically characterized in (4) (where $\vec{A} \to B$ abbreviates $A_1 \to \ldots A_n \to B$, and similiarly for $\vec{x}$).

$$\begin{array}{lll} (AR) & \vec{A} \to B \to \vec{C} \to D & \to \quad \vec{A} \to ((B \to D) \to D) \to \vec{C} \to D \\ & t & \to \quad \lambda \vec{x}_{\vec{A}} \lambda w_{(B\to D)\to D} \lambda \vec{y}_{\vec{C}}.w(\lambda z_B.t(\vec{x})(z)(\vec{y})) \end{array} \tag{4}$$

Directional realizations of this schema are *not* generally valid. In (5) and (6), one finds a case of derivational ambiguity which happens to be derivable, in the associative setting of **L**, for generalized quantifiers occupying peripheral positions in their scopal domain. The derivations represent two semantically distinct **L** proofs of the theorem

$$(np\backslash s)/np \Rightarrow ((s/(np\backslash s))\backslash s)/((s/np)\backslash s)$$

turning a simple first-order transitive verb into a third-order functor taking second-order generalized quantifier type arguments, encoding the subject wide scope reading (5) and object wide scope (6) reading, respectively. A change of the transitive verb type from $(np\backslash s)/np$ to $np\backslash(np\backslash s)$ (as would be required for verb-final languages such as German or Dutch), is enough to make the object wide scope reading underivable. What one would like to obtain in full generality is the possibility of having a generalized quantifier phrase at any np position, exerting its binding force at any s level of embedding.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\mathbf{tv}}{\mathbf{tv} : (np\backslash s)/np}\quad \overline{x_1 : np}}{\mathbf{tv}x_1 : np\backslash s}/\text{E}}{\overline{x_0 : np}}\backslash\text{E}}{\mathbf{tv}x_1x_0 : s}/\text{I}}{\lambda x_1.\mathbf{tv}x_1x_0 : s/np}\qquad \overline{x_2 : (s/np)\backslash s}}\backslash\text{E} \tag{5}$$

$$\dfrac{\overline{x_3 : s/(np\backslash s)}\qquad \dfrac{\dfrac{x_2(\lambda x_1.\mathbf{tv}x_1x_0) : s}{\lambda x_0.x_2(\lambda x_1.\mathbf{tv}x_1x_0) : np\backslash s}\backslash\text{I}}{x_3(\lambda x_0.x_2(\lambda x_1.\mathbf{tv}x_1x_0)) : s}/\text{E}}{\lambda x_3.x_3(\lambda x_0.x_2(\lambda x_1.\mathbf{tv}x_1x_0)) : (s/(np\backslash s))\backslash s}\backslash\text{I}$$

$$\dfrac{}{\lambda x_2\lambda x_3.x_3(\lambda x_0.x_2(\lambda x_1.\mathbf{tv}x_1x_0)) : ((s/(np\backslash s))\backslash s)/((s/np)\backslash s)}/\text{I}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\mathbf{tv}}{\mathbf{tv} : (np\backslash s)/np}\quad \overline{x_1 : np}}{\overline{x_0 : np}\qquad \mathbf{tv}x_1 : np\backslash s}\,/\mathrm{E}}{\mathbf{tv}x_1x_0 : s}\,\backslash\mathrm{E}}{\overline{x_2 : s/(np\backslash s)}\qquad \lambda x_0.\mathbf{tv}x_1x_0 : np\backslash s}\,\backslash\mathrm{I}}{x_2(\lambda x_0.\mathbf{tv}x_1x_0) : s}\,/\mathrm{E}}{\lambda x_1.x_2(\lambda x_0.\mathbf{tv}x_1x_0) : s/np}\,/\mathrm{I}\qquad \overline{x_3 : (s/np)\backslash s}}{x_3(\lambda x_1.x_2(\lambda x_0.\mathbf{tv}x_1x_0)) : s}\,\backslash\mathrm{E}}{\lambda x_2.x_3(\lambda x_1.x_2(\lambda x_0.\mathbf{tv}x_1x_0)) : (s/(np\backslash s))\backslash s}\,\backslash\mathrm{I}}{\lambda x_3\lambda x_2.x_3(\lambda x_1.x_2(\lambda x_0.\mathbf{tv}x_1x_0)) : ((s/(np\backslash s))\backslash s)/((s/np)\backslash s)}\,/\mathrm{I} \tag{6}$$

As an illustration for the FMG type-shifting approach, take the sentence 'Mary thinks someone left'. In (7) we list the necessary steps producing the wide scope reading for 'someone'. We give both the semantic shifts — abbreviating $A \to B$ as (AB) — and their directional counterpart. The (AR) transition for 'left', with the generalized quantifier variable x_2 in head position, is the critical one that cannot be obtained as a pure **NL** proof term. Combining the words (in their shifted types) by means of functional application produces the desired reading.

$$\begin{array}{ll}
\text{left} & np\backslash s \Rightarrow np\backslash((s/s)\backslash s) \\
& (et) \Rightarrow_{VR} (e((tt)t) \\
& \text{left} \Rightarrow \lambda x_1\lambda x_0.x_1(\text{left}(x_0)) \quad (= \text{left'}) \\
\\
& np\backslash((s/s)\backslash s) \Rightarrow (s/(np\backslash s))\backslash((s/s)\backslash s) \\
& (e((tt)t) \Rightarrow_{AR} ((et)t)((tt)t) \\
& \text{left'} \Rightarrow \lambda x_2\lambda x_0.x_2(\lambda x_1.x_0(\text{left}(x_1))) \\
\\
\text{thinks} & (np\backslash s)/s \Rightarrow (np\backslash s)/((s/s)\backslash s) \\
& (t(et)) \Rightarrow_{AR} ((tt)t)(et) \\
& \text{thinks} \Rightarrow \lambda x_2\lambda x_0.x_2(\lambda x_1.\text{thinks}(x_1)(x_0))
\end{array} \tag{7}$$

A CONNECTIVE FOR BINDING. The challenge for the deductive approach is to identify within the theoretical space of the substructural grammar logics a connective with the proper inferential capacity to render the critical AR cases (the cases beyond the reach of **(N)L**) derivable.

As a first approximation, (8) presents the connective $q(A, B, C)$ for 'in situ' binding proposed in [9]. Use of a formula $q(A, B, C)$ binds a variable x of type A, where the resource A is substituted for (takes the place of) $q(A, B, C)$ in the binding domain B. Using $q(A, B, C)$ turns the binding domain B into C. In the generalized quantifier case we have typing $q(np, s, s)$ where it happens that $B = C = s$. For the semantic term decoration of the rule of use $[qL]$, assume $t(q(A, B, C)) = (t(A) \to t(B)) \to t(C)$.

$$\dfrac{\Delta[x : A] \Rightarrow t : B \qquad \Gamma[y : C] \Rightarrow u : D}{\Gamma[\Delta[z : q(A, B, C)]] \Rightarrow u[z(\lambda x.t)/y] : D}\ (qL) \tag{8}$$

As an illustration of the use of the q connective, (9) gives direct cut-free proof search for 'John thinks someone left', with wide scope 'someone'. (Compare: the FMG strategy of (7).)

$$\dfrac{\dfrac{\dfrac{\dfrac{np \Rightarrow np \quad s \Rightarrow s}{np, np\backslash s \Rightarrow s}\,(\backslash L) \qquad \dfrac{np \Rightarrow np \quad s \Rightarrow s}{np, np\backslash s \Rightarrow s}\,(\backslash L)}{np, (np\backslash s)/s, \boxed{x : np}, np\backslash s \Rightarrow \boxed{u : s}}\,(/L) \qquad \boxed{y : s} \Rightarrow y : s}{np, (np\backslash s)/s, \boxed{\text{someone:}q(np, s, s)}, np\backslash s \Rightarrow \boxed{y[\text{someone}(\lambda x.u)/y] : s}}\,(qL) \tag{9}$$

$$u = \text{thinks}(\text{left}(x))(\text{j}),$$
$$y[\text{someone}(\lambda x.u)/y] = \text{someone}(\lambda x.\text{thinks}(\text{left}(x))(\text{j}))$$

Carpenter [1, 2] offers an in-depth discussion of the empirical range of the binding connective as compared with competing approaches to quantification (such as Cooper storage or LFG style type-driven interpretation), and he puts to use the binding connective in a treatment of plurals and collectivity. It may be worthwile to stress the different 'heuristic' qualities of the deductive approach and the rule-based type-shifting alternative. The type-shifting approach is specifically designed to handle the semantics of quantificational phenomena under minimal type assignment assumptions. The deductive approach introduces a *connective*, i.e. a fully general operation on types that cannot have a construction-specific limited range of application. Support for the generality of a connective for *in situ* binding can be found in the analyses of Pied-Piping ([14]), or *more ... than* comparative subdeletion ([5]).

The q connective, as presented above, raises a number of model-theoretic and proof-theoretic questions which we address in the following sections. With the $[qL]$ inference, we have given a rule of *use* — what about the rule of proof for the *in situ* binder? Can we give the interpretation of the q connective in terms of the general models for categorial logics, so that the proposed inference rules can be evaluated with respect to soundness and completeness? Is the q connective a primitive or a defined operator? Notice that the semantic term assignment for the $[qL]$ inference shows the interaction of two implications (the application and abstraction operations in $z(\lambda x.t)$) — could we decompose the q connective into more elementary logical constants, in a way that makes this implicational reasoning explicit?

3 Deconstructing q in an associative regime

In [13], one finds a multimodal analysis of the q connective which addresses these questions. The deconstruction of *in situ* binding presented in [13] assumes a multimodal setting with three families of multiplicatives: associative $\bullet_a$, non-associative $\bullet_n$ and a wrapping mode w which mediates between the two. The relevant structural postulates are presented below in axiomatic format.[1] The

1. The constant I has to be added to allow for cases where the *in situ* binder is peripheral in its domain. There are model-theoretic and proof-theoretic problems with a multiplicative identity which we ignore here, because the alternative to be presented in §4 does not make use of I.

interpretation of the multimodal type language and of the structural postulates with respect to the general categorial frame semantics or its groupoid specialization, and the transformation of the axiomatic presentation into Gentzen-style sequent format, are standard. We refer to [11] for a general discussion.

$$\begin{array}{ll} (I) & I \bullet_a A \longleftrightarrow A \longleftrightarrow A \bullet_a I \\ (A) & (A \bullet_a B) \bullet_a C \longleftrightarrow A \bullet_a (B \bullet_a C) \\ (WN) & (A \bullet_a B) \bullet_a C \longleftrightarrow (A \bullet_n C) \bullet_w B \end{array} \tag{10}$$

The interaction postulate WN is the crucial feature of this analysis: an associative configuration $(A \bullet_a B) \bullet_a C$ can be reanalysed as B in wrapping construction with a configuration $A \bullet_n C$. The non-associativity of $\bullet_n$ keeps track of the 'insertion point' for wrapping, the associativity of $\bullet_a$ makes any position in an associative configuration eligible as a potential locus for wrapping.

Given the postulates (10), the proposal of [13] is to define the connective $q(A, B, C)$ as $(B/_w A)\backslash_w C$. Compare the $[qL]$ rule of (8), repeated below for convencience without the term decoration, with the partial deduction for $(B/_w A)\backslash_w C$ in (12). (Here and below, we abuse notation a little to economize on brackets, and write the logical connectives $\bullet_i$ where we should have their structural counterparts $(\cdot, \cdot)^i$ in sequent antecedents.)

$$\frac{\Delta[A] \Rightarrow B \quad \Gamma[C] \Rightarrow D}{\Gamma[\Delta[q(A, B, C)]] \Rightarrow D}\ (qL) \tag{11}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{(\Delta \bullet_a A) \bullet_a \Delta' \Rightarrow B}{(\Delta \bullet_n \Delta') \bullet_w A \Rightarrow B}\ WN}{\Delta \bullet_n \Delta' \Rightarrow B/_w A}\ /_w R \qquad \Gamma[C] \Rightarrow D}{\Gamma[(\Delta \bullet_n \Delta') \bullet_w (B/_w A)\backslash_w C] \Rightarrow D}\ \backslash_w L}{\Gamma[(\Delta \bullet_a (B/_w A)\backslash_w C) \bullet_a \Delta'] \Rightarrow D}\ WN}{\Gamma[(\Delta \bullet_a q(A, B, C)) \bullet_a \Delta'] \Rightarrow D}\ (def) \tag{12}$$

As an illustration, (13) gives a derivation for the sentence 'John thinks someone left' which produces the wide-scope reading for 'someone', under the standard Curry-Howard term assignment regime. We delay unpacking of the lexical type assignments to facilitate reading.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\&c}{\text{John} \bullet_a (\text{thinks} \bullet_a (np \bullet_a \text{left})) \Rightarrow s}\ A}{((\text{John} \bullet_a \text{thinks}) \bullet_a np) \bullet_a \text{left} \Rightarrow s}\ WN}{((\text{John} \bullet_a \text{thinks}) \bullet_n \text{left}) \bullet_w np \Rightarrow s}\ \backslash_w R}{(\text{John} \bullet_a \text{thinks}) \bullet_n \text{left} \Rightarrow s/_w np} \qquad s \Rightarrow s}\ /_w L}{((\text{John} \bullet_a \text{thinks}) \bullet_n \text{left}) \bullet_w (s/_w np)\backslash_w s \Rightarrow s}\ (def)}{((\text{John} \bullet_a \text{thinks}) \bullet_n \text{left}) \bullet_w \text{someone} \Rightarrow s}\ WN}{((\text{John} \bullet_a \text{thinks}) \bullet_a \text{someone}) \bullet_a \text{left} \Rightarrow s}\ WN}{\text{John} \bullet_a (\text{thinks} \bullet_a (\text{someone} \bullet_a \text{left})) \Rightarrow s}\ A \tag{13}$$

On closer inspection, we see that the deconstruction of the connective $q(A, B, C)$ as $(B/_w A)\backslash_w C$ is not entirely faithful. The $[qL]$ rule (11) is formulated in a 'context-free' fashion: it does not impose any restrictions on the

environment of the $q(A, B, C)$ formula. In contrast, the decomposition of (12) in terms of (10) is parasitic on associativity: a $q(A, B, C)$ formula can only live in an associative environment, i.e. an environment where constituent structure sensitivity is lacking. In the absence of associativity assumptions for the default composition mode, the interaction postulate WN cannot realize the required scopal possibilities for a $q(A, B, C)$ formula. If one replaces $\bullet_a$ in the example (13) by a non-associative product, the derivation of the non-local scope reading fails. In the next section, we present an alternative decomposition of $q(A, B, C)$ which overcomes this limitation of the Morrill analysis.

4 A general multimodal analysis

The multimodal decomposition to be presented below is based on communication between a wrapping mode, which here we'll write as $\circ$, and a set of modes it interacts with. We first present 'in situ' binding from the perspective of the pure residuation logic **NL**, to make it clear that the analysis does not depend on any resource management assumptions about the environment of the $\circ$ mode, but solely on the postulates governing the interaction of wrapping with its structural context. The generalization to the case where $\circ$ interacts with a family of modes $\bullet_i$ is entirely straightforward. The required structural postulates are given in (14).

$$\begin{array}{lll} (P0) & \Diamond A \longleftrightarrow A \circ \mathbf{t} & (P0') \\ (P1) & (A \circ B) \bullet C \longleftrightarrow A \circ \langle l \rangle (B \bullet C) & (P1') \\ (P2) & A \bullet (B \circ C) \longleftrightarrow B \circ \langle r \rangle (A \bullet C) & (P2') \end{array} \tag{14}$$

For the binary wrapping mode, postulate $P0$ introduces a degenerate case $\Diamond A$, defined as the composition of A with a type constant $\mathbf{t}$. The constant functions as a structural placeholder. The interaction postulates $P1$ and $P2$ regulate the communication between $\bullet$ and $\circ$. The postulates are two-sided variants of the Mixed Associativity and Mixed Commutativity principles introduced in [12] for the analysis of head adjunction. In the left-to-right direction, the postulates $P1$ and $P2$ swap the dominance relation between $\circ$ and $\bullet$, for a $\circ$ configuration embedded as the left $(P1)$ or right $(P2)$ factor of a $\bullet$ configuration. In raising out $\circ$, the affected $\bullet$ configuration is marked with $\langle l \rangle$ or $\langle r \rangle$, respectively. The modal decoration makes the right-to-left direction of the interaction postulates deterministic — a crucial feature for keeping track of the structural position of the hypothetical A assumption bound by the $q(A, B, C)$ binder.

COMPILATION OF 'IN SITU' BINDING. Given the postulates (14), the connective $q(A, B, C)$ can be defined as $\Diamond(C/_w(\Box^{\downarrow} A\backslash_w B))$. The partial execution compilation of the $[qL]$ rule is presented in (15). Notice that there are no contextual restrictions on the occurrence of the q connective. Some notational sugaring is used to make things more legible: $\Diamond^l_r \Delta[\mathbf{t}]$ stands for a configuration $\Delta[\mathbf{t}]$ with a path of $\langle l \rangle, \langle r \rangle$ diamonds connecting the root to the placeholder $\mathbf{t}$. In the antecedent, we again abuse notation in writing the logical product connectives for their structural counterparts.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Delta[A] \Rightarrow B}{\Delta[\Diamond\Box^{\downarrow}A] \Rightarrow B}\,\Box^{\downarrow}L}{\Delta[\Box^{\downarrow}A\circ\mathbf{t}] \Rightarrow B}\,P0'}{\Box^{\downarrow}A\circ\Diamond^{l}_{r}\Delta[\mathbf{t}] \Rightarrow B}\,P1'/P2'}{\Diamond^{l}_{r}\Delta[\mathbf{t}] \Rightarrow \Box^{\downarrow}A\backslash_{w}B}\,\backslash_{w}R \qquad \Gamma[C]\Rightarrow D}{\Gamma[C/_{w}(\Box^{\downarrow}A\backslash_{w}B)\circ\Diamond^{l}_{r}\Delta[\mathbf{t}]]\Rightarrow D}\,/_{w}L}{\Gamma[\Delta[C/_{w}(\Box^{\downarrow}A\backslash_{w}B)\circ\mathbf{t}]]\Rightarrow D}\,P1/P2}{\Gamma[\Delta[\Diamond(C/_{w}(\Box^{\downarrow}A\backslash_{w}B))]]\Rightarrow D}\,P0}{\Gamma[\Delta[q(A,B,C)]]\Rightarrow D}\,(def) \tag{15}$$

NON-LOCAL SCOPE: SAMPLE DERIVATION. We illustrate the non-local scoping possibilities of $q(np, s, s)$ in (16), a derivation for the wide-scope reading of 'John thinks someone left'.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\&c}{\text{John}\bullet(\text{thinks}\bullet(np\bullet\text{left}))\Rightarrow s}}{\text{John}\bullet(\text{thinks}\bullet(\Diamond\Box^{\downarrow}np\bullet\text{left}))\Rightarrow s}\,\Box^{\downarrow}L}{\text{John}\bullet(\text{thinks}\bullet((\Box^{\downarrow}np\circ\mathbf{t})\bullet\text{left}))\Rightarrow s}\,P0'}{\text{John}\bullet(\text{thinks}\bullet(\Box^{\downarrow}np\circ\langle l\rangle(\mathbf{t}\bullet\text{left})))\Rightarrow s}\,P1'}{\text{John}\bullet(\Box^{\downarrow}np\circ\langle r\rangle(\text{thinks}\bullet\langle l\rangle(\mathbf{t}\bullet\text{left})))\Rightarrow s}\,P2'}{\Box^{\downarrow}np\circ\langle r\rangle(\text{John}\bullet\langle r\rangle(\text{thinks}\bullet\langle l\rangle(\mathbf{t}\bullet\text{left})))\Rightarrow s}\,P2'}{\langle r\rangle(\text{John}\bullet\langle r\rangle(\text{thinks}\bullet\langle l\rangle(\mathbf{t}\bullet\text{left})))\Rightarrow\Box^{\downarrow}np\backslash_{w}s}\,\backslash_{w}R \qquad s\Rightarrow s}{s/_{w}(\Box^{\downarrow}np\backslash_{w}s)\circ\langle r\rangle(\text{John}\bullet\langle r\rangle(\text{thinks}\bullet\langle l\rangle(\mathbf{t}\bullet\text{left})))\Rightarrow s}\,/_{w}L}{\text{John}\bullet(s/_{w}(\Box^{\downarrow}np\backslash_{w}s)\circ\langle r\rangle(\text{thinks}\bullet\langle l\rangle(\mathbf{t}\bullet\text{left})))\Rightarrow s}\,P2}{\text{John}\bullet(\text{thinks}\bullet(s/_{w}(\Box^{\downarrow}np\backslash_{w}s)\circ\langle l\rangle(\mathbf{t}\bullet\text{left})))\Rightarrow s}\,P2}{\text{John}\bullet(\text{thinks}\bullet((s/_{w}(\Box^{\downarrow}np\backslash_{w}s)\circ\mathbf{t})\bullet\text{left}))\Rightarrow s}\,P1}{\text{John}\bullet(\text{thinks}\bullet(\Diamond(s/_{w}(\Box^{\downarrow}np\backslash_{w}s))\bullet\text{left}))\Rightarrow s}\,P0 \tag{16}$$

In order to use the type $\Diamond(s/_{w}(\Box^{\downarrow}np\backslash_{w}s))$ for 'someone', we have to unpack the $\Diamond$ operator, thus introducing the wrapping mode $\circ$ required for the implications $/_{w}, \backslash_{w}$. This is effected by $P0$. The scope of the generalized quantifier expression is determined at the $[/_{w}L]$ step. This stage of the derivation is reached via a sequence of applications of the $P1$ and $P2$ structural postulates. The structure which the in-situ binder travels through to reach its scope-taking position is 'frozen' by the $\langle l\rangle, \langle r\rangle$ marking leading back to the place of origin, which is marked by $\mathbf{t}$. The hypothetical reasoning with respect to $\Box^{\downarrow}np$ assumption starting at $[\backslash_{w}R]$ traverses this path in the opposite direction, 'unfreezing' the $\langle l\rangle, \langle r\rangle$ markings as it passes them. At the point where the binding site is recovered via $P0'$, the basic residuation law $\Diamond\Box^{\downarrow}np \rightarrow np$ applies, cancelling the $\Diamond$ and $\Box^{\downarrow}$ operators.

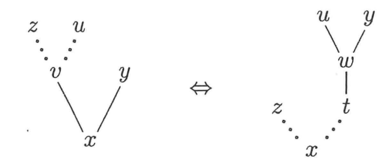
COMPLETENESS. We interpret the type language with respect to multimodal frames

$$F = \langle W, R^3_{\bullet}, R^3_{\circ}, R^2_{\Diamond}, R^2_{l}, R^2_{r}, *\rangle \tag{17}$$

We assume we have an object $* \in W$ for the interpretation of the placeholder constant, i.e. $v(\mathbf{t}) = \{*\}$. For the degenerate wrapping case $\Diamond$ we have the constraint (18) to the effect that $R^2_{\Diamond}$ is defined as $\{(x,y) \mid R^3_{\circ}xy*\}$.

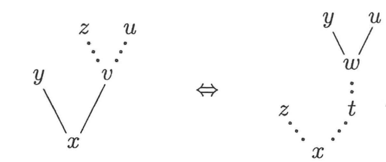
$$P0, P0' \qquad R^2_{\Diamond}xy \Leftrightarrow R^3_{\circ}xy* \tag{18}$$

The interaction postulates $P1$ and $P2$ have the required form for the Sahlqvist completeness result of [7]. One easily computes the corresponding frame constraints, using the Sahlqvist-van Benthem algorithm. (We write dotted lines for $R_{\circ}$ and R_{l}, firm lines for $R_{\bullet}$ and R_{r}.)



(19)

$(P1) \quad \exists v(R_{\bullet}xyv \;\&\; R_{\circ}vzu) \Leftrightarrow \exists t\exists w(R_{\circ}xzt \;\&\; R_{l}tw \;\&\; R_{\bullet}wyu) \quad (P1')$



$(P2) \quad \exists v(R_{\bullet}xvy \;\&\; R_{\circ}vzu) \Leftrightarrow \exists t\exists w(R_{\circ}xzt \;\&\; R_{r}tw \;\&\; R_{\bullet}wuy) \quad (P2')$

(20)

Notice that one could introduce $\bullet_{l}$, $\bullet_{r}$ as defined operators, for the compositions $\langle l\rangle(\cdot\bullet\cdot)$ and $\langle r\rangle(\cdot\bullet\cdot)$, respectively, and obtain (21) — saving some ink, but changing nothing in the setup of (17).

$$\begin{array}{lll} (p1) & (A\circ B)\bullet C \longleftrightarrow A\circ(B\bullet_{l} C) & (p1') \\ (p2) & A\bullet(B\circ C) \longleftrightarrow B\circ(A\bullet_{r} C) & (p2') \end{array} \tag{21}$$

We close with some remarks on the use of $q(A,B,C)$. The decomposition of 'in situ' binding as $\Diamond(s/_{w}(\Box^{\downarrow}np\backslash_{w}s))$ has been tested on the benchmark examples of scope ambiguity in [4]. In multiple quantifier interactions, the required scopal possibilities are produced — such as, for example, the eight readings for 'Fred thinks every schoolboy believes a mathematician wrote *Through the Looking-Glass*' displayed in (22). (We write GQ1 for 'every schoolboy' and GQ2 for 'a mathematician'.)

$$
\begin{array}{l}
(\text{GQ2}\ \lambda x_0.(\text{GQ1}\ \lambda x_1.((\text{THINK}\ ((\text{BELIEVE}\ ((\text{WRITE TLG})\ x_0))\ x_1))\ \text{FRED}))) \\
(\text{GQ1}\ \lambda x_0.(\text{GQ2}\ \lambda x_1.((\text{THINK}\ ((\text{BELIEVE}\ ((\text{WRITE TLG})\ x_1))\ x_0))\ \text{FRED}))) \\
(\text{GQ2}\ \lambda x_0.((\text{THINK}\ (\text{GQ1}\ \lambda x_1.((\text{BELIEVE}\ ((\text{WRITE TLG})\ x_0))\ x_1)))\ \text{FRED})) \\
((\text{THINK}\ (\text{GQ2}\ \lambda x_0.(\text{GQ1}\ \lambda x_1.((\text{BELIEVE}\ ((\text{WRITE TLG})\ x_0))\ x_1))))\ \text{FRED}) \\
(\text{GQ1}\ \lambda x_0.((\text{THINK}\ (\text{GQ2}\ \lambda x_1.((\text{BELIEVE}\ ((\text{WRITE TLG})\ x_1))\ x_0)))\ \text{FRED})) \\
((\text{THINK}\ (\text{GQ1}\ \lambda x_0.(\text{GQ2}\ \lambda x_1.((\text{BELIEVE}\ ((\text{WRITE TLG})\ x_1))\ x_0))))\ \text{FRED}) \\
(\text{GQ1}\ \lambda x_0.((\text{THINK}\ ((\text{BELIEVE}\ (\text{GQ2}\ \lambda x_1.((\text{WRITE TLG})\ x_1)))\ x_0))\ \text{FRED})) \\
((\text{THINK}\ (\text{GQ1}\ \lambda x_0.((\text{BELIEVE}\ (\text{GQ2}\ \lambda x_1.((\text{WRITE TLG})\ x_1)))\ x_0)))\ \text{FRED})
\end{array}
\tag{22}
$$

Notice that for the $\Diamond(s/_w(\Box^{\downarrow}np\backslash_w s))$ decomposition of the in situ binder, the following variant of the Geach law is derivable. This type transition would be appropriate to shift a regular sentence level generalized quantifier (e.g. 'someone') to a verb phrase level generalized quantifier (e.g. 'himself'). Under the analysis of §3, this type transition is underivable.

$$
\begin{array}{ll}
 & q(np,s,s) \rightarrow q(np,(np\backslash s),(np\backslash s)) \\
i.e. & \Diamond(s/_w(\Box^{\downarrow}np\backslash_w s)) \rightarrow \Diamond((np\backslash s)/_w(\Box^{\downarrow}np\backslash_w(np\backslash s)))
\end{array}
\tag{23}
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\&c}}{np \bullet (\Box^{\downarrow}np \circ \Box^{\downarrow}np\backslash_w(np\backslash s)) \Rightarrow s}\,P2'}{\Box^{\downarrow}np \circ \langle r\rangle(np \bullet (\Box^{\downarrow}np\backslash_w(np\backslash s)) \Rightarrow s}\,\backslash_w R}{\langle r\rangle(np \bullet (\Box^{\downarrow}np\backslash_w(np\backslash s)) \Rightarrow \Box^{\downarrow}np\backslash_w s} \quad s \Rightarrow s}{(s/_w(\Box^{\downarrow}np\backslash_w s)) \circ \langle r\rangle(np \bullet (\Box^{\downarrow}np\backslash_w(np\backslash s))) \Rightarrow s}\,/_w L}{np \bullet ((s/_w(\Box^{\downarrow}np\backslash_w s)) \circ (\Box^{\downarrow}np\backslash_w(np\backslash s))) \Rightarrow s}\,P2}{s/_w(\Box^{\downarrow}np\backslash_w s) \Rightarrow (np\backslash s)/_w(\Box^{\downarrow}np\backslash_w(np\backslash s))}\,/_w R, \backslash R}{\Diamond(s/_w(\Box^{\downarrow}np\backslash_w s)) \Rightarrow \Diamond((np\backslash s)/_w(\Box^{\downarrow}np\backslash_w(np\backslash s)))}\,\Diamond R
\tag{24}
$$

Notice also that for the deconstruction of q proposed here, the 'modal' form of the lifting law $A \rightarrow B/(A\backslash B)$ is (25a), not (25b). For conjunction of proper names with generalized quantifier noun phrases, one uses any of the 'regular' lifted np types, cf. (25c) and (25d).

$$
\begin{array}{ll}
(a) & \vdash \Diamond\Box^{\downarrow}A \rightarrow \Diamond(B/_w(\Box^{\downarrow}A\backslash_w B)) \\
(b) & \not\vdash A \rightarrow \Diamond(B/_w(\Box^{\downarrow}A\backslash_w B)) \\
(c) & \vdash \Diamond(s/_w(\Box^{\downarrow}np\backslash_w s)) \rightarrow s/(np\backslash s) \\
(d) & \vdash np \rightarrow s/(np\backslash s)
\end{array}
\tag{25}
$$

As remarked above, one can generalize $P1, P2$ to multimodal interaction principles by replacing $\bullet$ by $\bullet_i$, and specifying the set of modes i which the wrapping mode $\circ$ interacts with. One can also differentiate the wrapping mode itself, and distinguish wrapping variants that would interact with different contextual environments. Such distinctions would make it possible to lexically discriminate between generalized quantifiers with different scopal possibilities. We leave further elaboration for another occasion.

Acknowledgements

I thank Natasha Kurtonina for comments on an earlier draft, and Richard Moot for spotting a flaw in the submitted abstract, and for suggesting a remedy.

References

[1] Carpenter, B. (1994), 'Quantification and scoping: a deductive account'. *Proceedings 13th West Coast Conference on Formal Linguistics*. San Diego.

[2] Carpenter, B. (1996), *Type-Logical Semantics*. MIT Press, Cambridge, MA.

[3] Emms, M. (1990) 'Polymorphic quantifiers'. Proceedings 7th Amsterdam Colloquium. ILLC, Amsterdam.

[4] Hendriks, H. (1993) *Studied Flexibility*. PhD Dissertation, ILLC, Amsterdam.

[5] Hendriks, P. (1995a) *Comparatives in Categorial Grammar*. Ph.D. Thesis, Groningen.

[6] Hendriks, P. (1995b) 'Ellipsis and multimodal categorial type logic'. In Morrill (ed) *Proceedings Formal Grammar Conference*. Barcelona.

[7] Kurtonina, N. (1995) *Frames and Labels. A Modal Analysis of Categorial Inference*. PhD Dissertation, OTS Utrecht, ILLC Amsterdam.

[8] Kurtonina, N. & M. Moortgat (1995), 'Structural Control'. In M. Moortgat (ed.) *Logics of Structured Resources*. Esprit BRA 6852 Dyana-2 Deliverable R1.1.C. (To appear in P. Blackburn & M. de Rijke (eds.) *Logic, Structures and Syntax*. Reidel, Dordrecht.)

[9] Moortgat, M. (1991), 'Generalized quantification and discontinuous type constructors'. (To appear in W. Sijtsma & A. van Horck (eds.) *Discontinuous Constituency*, De Gruyter, Berlin.)

[10] Moortgat, M. (1995), 'Multimodal linguistic inference'. *Bulletin of the IGPL* **3**(2,3). Special issue on Deduction and Language (ed. R. Kempson), 371–401. (To appear in *Journal of Logic, Language, and Information*).

[11] Moortgat, M. (1996), 'Categorial Type Logics'. Chapter Two in J. van Benthem & A. ter Meulen (eds.) *Handbook of Logic and Language*. Elsevier, to appear.

[12] Moortgat, M. and R.T. Oehrle (1994), 'Adjacency, dependency and order'. In P. Dekker and M. Stokhof (eds.) *Proceedings Ninth Amsterdam Colloquium*, ILLC, Amsterdam, 447–466.

[13] Morrill, G. (1994) *Type Logical Grammar*. Kluwer, Dordrecht.

[14] Morrill, G. (1995), 'Discontinuity in Categorial Grammar'. *Linguistics & Philosophy* **18**, 175–219.

The Semantics of Paths and Spatial Orientations

Seungho Nam
Seoul National University

This paper aims to provide a formal semantic analysis of spatial expressions in English, so to build up a general logic of space to give compositional interpretations to sentences containing locative prepositional phrases (PPs). The logic of space proposed here is an enriched version of traditional topology by Tarski (1927/1956), and it introduces two fundamental notions, *path* and *orientation*, and defines their structures. Thus we extend the ontological domain with spatial objects called *regions*.

Nam (1995a) proposes the following typology of locative PPs, where belongs four classes of locatives. This paper, however, mainly deals with two of them: (1b) *Symmetric* locatives, and (1d) *directional* locatives.

(1) Four Classes of Locative PPs
   a. *Topological Invariants* — PPs with *at, in, on*
   b. *Symmetric Locatives*
      — PPs with *across, through, over, past, around*
   c. *Orientational Locatives*
      — PPs with *in front of/behind, to the left/right of, above/below*
   d. *Directional Locatives* — PPs with *to, from, into, out of, towards*

First, before getting into logical technicals, in section 1 we illustrates some semantic facts revealing two notions underlying the semantics of locatives —*symmetry* and *homogeneity*.

1. Symmetry and Homogeneity in Locatives

1.1. Symmetry-1

Locative PPs combine with a predicate and denote a predicate modifier. Crucially they are interpreted as determining a spatial property for an argument of the predicate. For example, in (2a), *in front of the house* determines a property 'being in front of the house' for the object argument 'Mary', so (2a) entails (2b).

(2) a. *John saw Mary in front of the house*
   b. ⊨ *Mary was in front of the house*

(3) a. *John spied on Mary from the control tower*
   b. ⊨ *John was in the control tower*
   c. ⊨ *Mary was not in the control tower*

In (3a), the PP *from the control tower* determines a property 'being in the control tower' for both the subject and the object argument. The property specifies the location of the subject argument 'John,' and its complement the location of the object 'Mary.' Thus (3a) entails both (3b) and (3c). Now consider the following with a symmetric locative.

(4) a. *John saw Mary through the window*
   b. ⊨ *John and Mary were on the opposite sides of the window*

(5) a. *John touched Mary across the table*
b. ⊨ *John and Mary were on the opposite sides of the table*

The PP in (4a), unlike those in (2) and (3), does not determine a unary property but a binary relation 'being on the opposite sides of the window,' and this relation specifies the spatial relation between the subject and the object arguments. Thus (4a) entails (4b), and the same is true of the entailment in (5a) and (5b). The following data also show that the symmetric PPs determine a symmetric binary relation.

(6) a. *John shouted at Mary over the fence*
b. *John heard the baby cry around the corner*

But notice that the binary relations induced by symmetric locatives are symmetric and are not reducible in terms of unary relations (properties): That is, they are not paraphrasable as a boolean compound of unary properties, so inherently binary.

1.2. Symmetry-2

Consider the following entailment pattern for another symmetry effect of symmetric locatives:

(7) a. *John walked across the street, and came back*
b. ⊨ *John walked across the street twice*

(8) a. *John ran around the corner, and came back*
b. ⊨ *John ran around the corner twice*

(7a) and (8a) entail (7b) and (8b), respectively. This entailment pattern is due to the symmetric nature of the paths determined by the locative PPs, *across the street* in (7a) and *around the corner* in (8a). In other words, symmetric locatives do not refer to the direction of movement. So for (7a) to be true, it does not matter which side of the street John started from, but he only had to cross the street walking. Other symmetric PPs with *over, past,* and *through* follow the same entailment pattern. It is not surprising, however, non-symmetric locatives do not follow the entailment pattern.

(9) a. *John walked in/into the room, and came back*
b. ⊭ *John walked in/into the room twice*

(10) a. *John drove to Amsterdam, and came back*
b. ⊭ *John drove to Amsterdam twice*

(11) a. *John ran in front of City Hall, and came back*
b. ⊭ *John ran in front of City Hall twice*

The PPs in (9) and (10) are directional locatives, and the PP *in front of the City Hall* in (11) is an orientational one. (a)-sentences in (9-11) do not entail (b)-sentences.

1.3. Symmetry-3

Some locatives require a *locative perspective* (or *locative point of view*) to get a proper interpretation. Here by *locative perspective*, we refer to a spatial setting which determines a relevant spatial orientation in question. For example, *across the street* in (12a) does not tell us which side of the street the subject argument 'John' is located. So we need another point to

determine on which side of the street the PP locates 'John.' The locative perspective may be supplied by the utterance context, so the utterance place 'here' often serves as the perspective point, but not always. (12b) shows that the locative perspective can be overtly expressed by a *from*-phrase, which by no means denotes a starting/source point of a movement. When the same PP modifies a motion verb like *ran* in (12c), however, *from*-phrase denotes a source point of the movement path. Thus, (12c) does not require a locative perspective.

(12) a. *John is sitting across the street*
b. *John is sitting across the street from here*
c. *John ran across the street (from here)*

(13) a. *The village is through the forest from here*
b. *The post office is around the corner from here*
c. *The boys were playing over the wall*
d. *The library is past the book store*

The sentences in (13) above give more examples with symmetric locatives whose interpretation requires a locative perspective. Non-symmetric locatives in (14) below do not allow a locative perspective. Thus, (14b) and (15b) are ungrammatical.

(14) a. *John is sitting in front of the car*
b. **John is sitting in front of the car from here*

(15) a. *John is sitting inside the room*
b. **John is sitting inside the room from here*

1.4. Symmetry-4

When a frequency adverbial occurs with a symmetric locative as in (16), we might get an ambiguity: One reading of (16a) is an *event-counting* reading, where John's jogging happens twice everyday; and the other reading is a *path-counting* reading, where the total length of John's jogging for a day amounts to two turns around the park. (16b) shows that separate frequency adverbials can occur to quantify over either events or paths. Thus *three times* refers to the number of events and *twice* the number of paths. If a frequency adverbial refers to a fraction as in (17), we do not get an event-counting reading.

(16) a. *John jogs around the park twice everyday*
b. *John jogs around the park three times twice a day*

(17) a. *John jogs around the park twice and half everyday*
b. *John swam across the pool twice and half everyday*

(18) a. *John drove down into the city twice everyday*
b. **John drove down into the city twice and half everyday*

We note here that path-counting reading is not available for non-symmetric locatives. When a frequency adverbial occurs with a non-symmetric locative, it always quantifies over events. (18a) is not ambiguous but has an event-counting reading only.

1.5. Homogeneity: Telic/Atelic Aspect

Directional prepositions *to* and *towards* both refer to the direction of a

movement, but the reference object of *towards* does not refer to the goal of the movement, whereas that of *to* determines the goal. Thus, (19a) but not (19b) entails (19c).

(19) a. *John ran to the post office*
b. *John ran towards the post office*
c. *John was at the post office*

Hinrichs (1985:204, 1986:349) observes the aspectual difference between *to* and *towards* in the following examples.

(20) a. *Fangs slithered to the rock*
b. *Fangs slithered toward the rock*
c. *John walked to the library in an hour*
d. **John walked toward the library in an hour*
e. *It took Fangs ten minutes to slither to the rock*
f. **It took Fangs ten minutes to slither toward the rock*

Hinrichs, in Vendler's (1967) terminology, classifies (20a) and (20b) as *accomplishment* and *activity*, respectively, and so the test results in (20c-f). In section 4, we will give lexical semantics of *to* and *towards*, and account for the aspectual difference by semantically characterizing a class of locatives which denote a *homogeneous* path.

2. The Natural Logic of Space

2.1. Mereological Primitives

Our logic of space is based on the primitive notion of *region* which enables us to treat locative PPs as denoting predicate modifiers represented in terms of properties and relations over regions. Locative PPs now supply such relevant properties and relations for its compositional interpretation. Based on these ontological primitives, we define other notions like *path* and *orientation.*

We start with a mereological space Σ structured by the primitive *part-to-whole* relation $\subseteq$. The elements in Σ are called *regions*, so Σ is the set of regions.

(21) The space $<\Sigma, \subseteq, \emptyset$, Between, Nearer$>$:
a. Σ: The set of *regions*
b. $\emptyset$: The special element (the *empty region*)
For any region $A \in \Sigma$, $\emptyset \subseteq A$

c. $\subseteq$: Binary *part-to-whole* relation in Σ
•Reflexive, transitive, and antisymmetric: Partial order

d. Between: Ternary *betweenness* relation in Σ
•Transitive:[1] If Between(X, Y, Z) and Between(Z, Y, U),
then Between(X, Y, U)

[1] The transitivity axioms for Between and Nearer are given in the same format: That is, for a ternary relation R, R is transitive iff if R(X, Y, Z) and R(Z, Y, U), then R(X, Y, U). But the transitivity axiom could be given in other forms: For example, Tarski's (1959) axiom is the following:
Between(Y, X, Z) and Between(U, Y, Z), then Between(Y, X, U).

•Symmetric on 2nd-3rd arguments:
Between(X, Y, Z), then Between(X, Z, Y)

e. Nearer: Ternary *relative nearness* relation in Σ
•Irreflexive: $\neg$Nearer(X, Y, X) & $\neg$Nearer(Y, X, X)
•Transitive: If Nearer(X, Y, Z) & Nearer(Z, Y, U),
then Nearer(X, Y, U)
•Asymmetric on 1st-3rd arguments:
If Nearer(X, Y, Z), then $\neg$Nearer(Z, Y, X)

(21) introduces two primitive ternary relations among regions, which impose geometric structures on the space Σ. First, the betweenness relation (Between) is given as transitive and symmetric on its second and third arguments, and Between(X,Y,Z) is intended to mean "X lies between Y and Z." The other ternary relation, Nearer(X,Y,Z) is intended to mean "X is nearer to Y than Z is," and it is given as irreflexive, transitive, and asymmetric on its first and third arguments.[2]

Tarski's (1959) axiom system for Elementary Geometry introduces a ternary primitive relation, *betweenness*, and a quaternary primitive relation, *equidistance* (AB=CD: "A is as distant from B as C is from D."). Our formalism with the relative nearness is richer than Tarski's, since the relative nearness relation can easily define the equidistance relation. Robinson (1959) shows that it is impossible for one or more binary relations to serve as the primitive notions in Euclidean geometry. Thus the ternary nature of spatial relations reveal that the space is more complex than the temporal domain which is usually characterized by the binary relations, *precedence* and *overlap* relations (e.g., Kamp 1979; van Benthem 1983).

2.2. Path Structure

Meanings of the spatial expressions in English can be categorized into two groups: (i) *movement-directional*, and (ii) *stative-locational*. Movement-directional meaning emerges when the sentence describes change of location, and stative-locational meaning is identified when an event takes place in a stative context without changing location. Our logic accomodates this contrast in terms of *path* and *orientation.*

Path is one of the basic concepts discussed in the literature on spatial language and it is claimed to be a crucial notion in perception/cognition of movement or journey, and it is one of the main cognitively motivated devices for representing changes of location (see Miller and Johnson-Laird (1976), Cresswell (1978) and Jackendoff (1983, 1990) among others). We define a path as a sequence of regions. The notion of path we are interested in here is not of course a physical one but an abstract one. The same physical path then can be represented as different abstract paths: If John flew from Seoul to Hong Kong and then to Amsterdam, John's path can be

[2] Euclidean primitives--*point* and *line*--can be defined in terms of *atomic regions* and *betweenness* relation as follows: (Def.-1) A non-empty region A is an *atomic region* if there is no region $B \neq \emptyset$ such that $B \subseteq A$ but $B \neq A$; (Def.-2) Let A and B be any distinct atomic regions in Σ. Define L_{AB}, the *line determined by A and B*: $L_{AB} =_{df.} \{X \in \Sigma \mid$ X is an atomic region and either X=A or A=B or Between(A,X,B), or Between(X,A,B), or Between(B,A,X)$\}$.

represented either as <**Seoul**, **Hong-Kong**, **Amsterdam**> or as <**Seoul**, **Amsterdam**>. The abstract notion of path now lacks the continuity of physical movement (or geometric lines). This is what renders our semantics of space sufficiently elegant as to interpret spatial expressions.

We define *Path Structure* as a set of paths which are partially ordered by the *subpath* relation. Paths are defined as sequences of regions, which are time-free. A sequence is a function with a domain of natural numbers from zero through some k, represented as $[0, k]$ in (22).

(22) *Path Structure*: $\langle\prod(\Sigma), \subseteq, +\rangle$:

a. $\prod(\Sigma)$: The set of paths in a local space Σ.

b. A path π is a sequence of regions, i.e., $\pi\in[[0, k]\rightarrow\Sigma]$ for some $k\in\mathbb{N}$, where $[0, k] = \{n\in\mathbb{N} \mid 0\leq n\leq k\}$, and satisfies the following: $\forall\pi\in\prod(\Sigma)$, and $\forall i\in\text{Domain}(\pi)$, $\pi(i-1) \prec \pi(i) \prec \pi(i+1)$.

c. $\subseteq$: Binary relation (*subpath*) between paths defined by:
Let π and π' be paths, then π' *is a subpath of* π, $\pi'\subseteq\pi$, if
(i) $\text{Domain}(\pi')\subseteq \text{Domain}(\pi)$ and $\text{Range}(\pi') \subseteq \text{Range}(\pi)$, and
(ii) there is some $i\in\text{Domain}(\pi)$ such that $\pi'(0) = \pi(i)$ and for all $j\in\text{Domain}(\pi')$, $\pi'(j) = \pi(i+j)$.

d. $+$: *Concatenation* function in $[\prod(\Sigma)\times\prod(\Sigma)\rightarrow\prod(\Sigma)]$ defined as:
Let π and π' be arbitrary paths with $\text{Domain}(\pi) = [0, n]$ and $\text{Domain}(\pi') = [0, m]$. The *concatenation* of π and π', $\pi+\pi'$, is defined by:

$$\pi+\pi'(k) = \begin{cases} \pi(k) & \text{if } 0 \leq k < n \\ \pi(k) = \pi'(0) & \text{if } k = n \\ \pi'(k-n) & \text{if } n < k \leq n+m \end{cases}$$

(22b) imposes a general condition on the path structure which is linguistically motivated by the following sentences:

(23) a. **John drove from Amsterdam to the Netherlands*
b. **John drove from the Netherlands to Amsterdam*

The locative PPs in (23a) and (23b) make the sentences meaningless since they fail to refer to a legitimate path. Thus we want to rule out paths such that some region in the path-sequence is included in the next region of the path, or vice versa. There is no path in $\prod(\Sigma)$ like the following: <**Amsterdam, Netherlands**> or <**Netherlands, Amsterdam**>.

From the definition of *subpath relation*, we can prove the relation is reflexive, transitive, and antisymmetric. Also notice that the *concatenation* function is not a total function, i.e., for some path π and π', $(\pi+\pi')$ is not defined. This is due to the condition given in the definition, i.e., the last region of π and the first region of π' have to coincide, i.e., $\pi(k) = \pi'(0)$. In other words, concatenation of π and π' exists iff the goal of π and the source of π' are identical (*goal* and *source* are defined below).

(24) Definitions:
Let π be a path with $\text{Domain}(\pi) = [0, k]$, then
the *goal* of π, π_g, is $\pi(k)$; and the *source* of π, π_s, is $\pi(0)$.

Now we define a very special relation between paths—*converse relation*. For all paths π, we have a path which reverses the ordering of π. We

define:

(25) Definition:
Let π be a path with $\text{Domain}(\pi) = [0, k]$, then
π^{-1}, the *converse* or π (or "π-converse") is defined by:
$\text{Domain}(\pi^{-1}) = [0, k]$, and
for all $i\in\text{Domain}(\pi^{-1})$, $\pi^{-1}(i) = \pi(k-i)$.

By the definition, $\pi^{-1}(0) = \pi(k)$, $\pi^{-1}(k) = \pi(0)$, i.e., the source of π^{-1} is the goal of π and the goal of π^{-1} is the source of π. We crucially rely on the notion of *path-converse* for the semantics of symmetric locatives and some special adverbs/verbs like *back* and *return*.

We note two advantages of our semantics of path: (i) our notion of path is not temporal, so paths are introduced as purely spatial entities; (ii) it is flexible enough to accommodate cyclic paths that allow some regions can occur more than once in a single path.

The intuitive notion of *path* involves a movement of an object. Now to represent the notion of movement through a path, we introduce a predicate TRAV which is a ternary relation in $\mathbb{E}\times\prod(\Sigma)\times\mathbb{T}$, where $\mathbb{E}$ is the universe of individuals, $\prod(\Sigma)$ the set of paths in the local space Σ, and $\mathbb{T}$ the set of time intervals. Informally, TRAV(x, π, T) means 'x traverses the path π during the interval T.' Here π is a sequence of regions with its domain in natural numbers, $\mathbb{N}$, and T a linearly ordered set of time points. In order to define this predicate formally, we use the function $\text{®}\in[(\mathbb{E}\times\mathbb{T})\rightarrow\Sigma]$ which assigns a unique region to each individual object at an interval. Thus, for some object x, and a time interval T, ®(x)(T) denotes the region which x occupies during the interval T. Now formally,

(26) Definition:
TRAV(x, π, T) is True iff there is an "order-preserving" map μ from $\text{Range}(\pi)$ to T such that for all $i\in\text{Domain}(\pi)$, $\text{INTR}(\text{®}(x)(\mu(\pi(i)), \pi(i))$.

We take an interval T as a linearly ordered set of time points, but the domain of μ is not an ordered set but a sequence (functions from $K\subseteq\mathbb{N}$ into Σ). Thus we use the term "order-preserving" in a special sense defined as follows:

(27) Definition:
For a path π and an interval T, a function μ from $\text{Range}(\pi) = \{\pi(i) \mid i\in\text{Domain}(\pi)\}$ to T is *order-preserving* iff
for all $i, j\in\text{Domain}(\pi)$, if $i \leq j$, then $\mu(\pi(i)) \leq_T \mu(\pi(j))$ and $\mu(\pi_s) = \mathbf{0}_T$ and $\mu(\pi_g) = \mathbf{1}_T$.

An interval T is ordered by the precedence relation ($\leq_T$), and the least element of T, $\mathbf{0}_T$, refers to the initial point of the interval T, and the greatest element $\mathbf{1}_T$ refers to the terminal point of T. The TRAV relation will be used to interpret sentences referring to a path and a movement. For example, *John ran into the house* will be interpreted to be true iff 'John ran' and 'John traversed the path π, i.e., TRAV(**john**,π,T), such that the source of the path is outside the house and the goal is inside the house'.

2.2.1 Theorems on Path Structure and Relations

Given the *Path Structure* and the relation TRAV, we can account for some

general entailment patterns induced by spatial expressions. (28) and (29) illustrate an entailment pattern based on the part-to-whole relation between regions.

(28) a. *John came from Seoul* entails
b. ⊨ *John came from Korea* (given Seoul is in Korea)

(29) a. *John drove from Paris to Amsterdam* entails
b. ⊨ *John drove from Paris to the Netherlands* and
c. ⊨ *John drove from France to the Netherlands*

The PPs in (28) and (29) determine a path referring to a source or a goal region, and part-to-whole relation holds between Seoul and Korea, etc. Thus (28a) and (29a) entail (28b) and (29b-c), respectively. The following theorem accounts for the entailment.

(30) Theorem-1:
For all paths π and π', and intervals T, TRAV(x, π, T) implies TRAV(x, π', T), if π and π' satisfy the following:
(i) π is a path such that Domain(π) = [0, k] and
(ii) for some $0 \le i \le j \le k$, and a region $R \in \Sigma$,
$\cup$Range($\pi_{[i,j]}$) $\subseteq$ R, and π' is defined as follows:
Domain(π') = [0, k–(j–i)], and

$$\pi'(n) = \begin{cases} \pi(n) & \text{if } 0 \le n < i \\ R & \text{if } n = i \\ \pi(n-(j-i)+1) & \text{if } i < n \le k \end{cases}$$

(31) Notation:
For π a path with the domain of [0, k], and for some $0 \le i \le j \le k$, $\pi_{[i,j]}$ is a subpath of π defined by:
Domain($\pi_{[i,j]}$) = {0,...,(j-i)}, and
$\pi_{[i,j]}(n) = \pi(n+i)$ for all $n \in$ Domain($\pi_{[i,j]}$)

Let us now consider (32) and (33) below where the entailment patterns involve the notion of "subpath": (32a) entails (32b-c), and (33a) entails (33b-e).

(32) a. *John flew from Los Angeles to San Diego, and then to Las Vegas*
b. ⊨ *John flew from Los Angeles to Las Vegas* and
c. ⊨ *John flew from San Diego to Las Vegas*

(33) a. *John drove through the forest from here to the village* entails
b. ⊨ *John drove from here to the village*
c. ⊨ *John drove from here (through the forest)*
d. ⊨ *John drove (through the forest) to the village* and
e. ⊨ *John drove through the forest*

The intuition is, if an object traversed a path π, then it must have traversed the subpaths of π. The following theorem accounts for the entailment pattern.

(34) Theorem-2:
Let π and π' be paths such that π' is a subpath of π.
If TRAV(x, π, T), then TRAV(x, π', T') for some T'
a subinterval of T.

2.3 Orientation Structure

Natural language expressions refer to spatial orientations to locate some object in the space, so for example, the PP in (35) refers to the *front-orientation* of the car to locate the subject argument 'John.'

(35) *John is sitting in front of the car*

Thus the sentence is true only if John's region is on the front-orientation of the car. Roughly, the front-orientation of a car can be characterized as a half axis moving out from the car in the direction to its front part. That is, we take orientations as spatial objects like rays which have a designated region (=Origin) and a direction. Rays, in their geometric sense, are collections of atomic regions. *Orientation Structure* is defined as the set of orientations with the *containment* relation ($\subseteq_R$) between them.

(36) Orientation Structure: $(\mathbf{R}(\Sigma), \subseteq_R>$:
a. $\mathbf{R}(\Sigma)$: The set of orientations in a local space Σ
b. $\subseteq_R$: Binary relation (containment) between orientations.
c. An orientation $\rho \in \mathbf{R}(\Sigma)$ is a linearly ordered set (ρ, <)
of atomic regions such that
(i) there is a unique least element (Origin(ρ)),
(ii) there is an atomic region $X \in \rho$ such that for all atomic regions $Y \in \Sigma$, $Y \in \rho$ iff either Between (Y, Origin(ρ), X) or Between (X, Origin(ρ), Y), and
(iii) for all atomic regions X, $Y \in \rho$,
X < Y iff Nearer (X, Origin(ρ), Y)

The linear order for an orientation is intended to be the relative distance relation from the origin of the orientation, as defined in (36c.iii). A linear order is total, irreflexive, asymmetric, and transitive. The betweenness condition given in (36c.ii) guarantees that for any orientation ρ, there is a line such that all the atomic regions in ρ belong to the line. Orientations will be used to interpret symmetric locatives as well as orientational locatives.

Object internal properties determine their intrinsic orientations, e.g., front/back, top/bottom, left/right, and in/out orientations. We claim here that such objects get their intrinsic orientations due to their different parts. In other words, if an object has inherent front and back parts, it can be assigned front/back orientations determined by them. For example, a car has an inherent front part, no matter how it is determined (whether it is determined by the normal direction of movement or by a formal characteristics of its front part), so we can assign front or back orientation to it. In other words, we think of *intrinsic orientation* as a derived concept from parts of objects. The following illustrate different primitive functions partitioning an object into different parts. For instance, $\mathbf{P}_{front}$(the car) refers to the front part of the car, and so on. These functions are partial since objects may lack some (or all) of the intrinsic orientations.

(37) $\mathbf{P}_{front}(x)$, $\mathbf{P}_{back}(x)$, $\mathbf{P}_{top}(x)$, $\mathbf{P}_{bottom}(x)$, $\mathbf{P}_{left}(x)$, $\mathbf{P}_{right}(x)$, $\mathbf{P}_{inside}(x)$, $\mathbf{P}_{outside}(x)$, $\mathbf{P}_{center}(x)$

Now we define orientations as determining a ray with an origin and a direction in terms of parts of objects (equivalently, a linearly ordered set of minimal regions with a least element), and such orientations are given as a function from pairs of an object and a time interval to rays, i.e.,

$[\mathbb{E}\times\mathbb{T}]\rightarrow\mathbf{R}(\Sigma)$, where $\mathbb{E}$ is the universe of individuals, $\mathbb{T}$ a set of intervals, and $\mathbf{R}(\Sigma)$ the set of orientations in Σ.

(38) $\Omega_{\text{front}}(x,t)$ is an orientation ρ such that
Origin(ρ) = ®($\mathbf{P}_{\text{center}}(x)$,t) and
ρ intersects ®($\mathbf{P}_{\text{front}}(x)$,t), i.e., $\rho\cap$®($\mathbf{P}_{\text{front}}(x)$,t) $\neq\emptyset$.

(39) $\Omega_{\text{back}}(x,t)$, $\Omega_{\text{up}}(x,t)$, $\Omega_{\text{down}}(x,t)$, $\Omega_{\text{left}}(x,t)$, $\Omega_{\text{right}}(x,t)$

Each function assigns an orientation to an object at an instance, and the orientation has its origin at the center of the object and intersects its relevant part. *In front of the car*, for example, involves the front orientation of 'the car', $\Omega_{\text{front}}(\mathbf{the\ car},t)$, which is a linearly ordered set of minimal regions with its origin at the center of 'the car' and directed to its front side. The following illustrate some objects with their intrinsic orientations:

(40) *man, car*: Top/Bottom, Front/Back, Right/Left, In/Out
telephone: Top/Bottom, Front/Back, In/Out
vase, boulder: Top/Bottom, In/Out
tree: Top/Bottom
rocket: Front/Back
box, ball: In/Out

Finally we define a binary relation INTR ("intersection") between regions and extend it to hold between a region and an orientation.

(41) Definitions
a. A region A is called an *intersection* of all elements of a set R of regions ($\cap$R) if A is non-empty part of each element of R and if there is no region B such that B is a part of each element of R and A is a proper part of B.
b. Regions A and B *intersect* each other, INTR(A, B), if there is an intersection of {A, B}.
c. For *X* a region and ρ a set of regions, INTR(*X*, ρ), *X intersects* ρ, iff there is a region $A\in\rho$ such that $A\subseteq X$.

(41b) defines INTR relation to hold between regions, and (41c) extends the relation to hold between a region and a set of regions. Thus a region may bear the INTR relation to either an orientation or a path.

2.3. Two General Interpretative Rules

Based on the formal structures of paths and orientations, we now interpret locatives with a couple of rules general enough to fit the different locative PPs. Following the general dichotomy of movement vs. stative readings of locatives, we interpret locatives as denoting paths or orientations. A path is associated with a movement reading induced by a motion verb, and an orientation is associated with a stative reading induced by a stative verb. The first rule in (42) is for interpreting sentences with a motion verb and a locative PP, which naturally induces a path that an object traverses.

(42) Semantic Rule–1:
For *m* a one-place motion verb, and *f* an extensional locative modifier, interpret the VP *m+f*, the concatenation of *m* and *f* by:
f(**m**)(x) iff **m**(x) and TRAV(x, π_f, T),
where π_f is a path determined by *f*.

Different locatives determine different paths. For example, *into the library* refers to a path whose source is a region outside the library, and whose goal is a region inside the library, and *over the fence* refers to a path whose source and goal are on the opposite sides of the fence, etc. Thus, *John jumped over the fence* is true if and only if 'John jumped during a past interval T and John traversed the path from one side of the fence to the other during T.'

The second rule in (43) is for interpreting sentences with a stative verb and a locative PP, so their interpretations do not involve a change of location.

(43) Semantic Rule–2:
For *s* a one-place stative verb, and *f* an extensional locative modifier, interpret the VP *s+f* by:
f(**s**)(x) iff **s**(x) and INTR(®(x), ρ_f)
where ρ_f is a an orientation or a region determined by *f*.

Orientational locatives like *in front of the house* determines an orientation (front/back, left/right, top/bottom), and sometimes the orientation is an intrinsic one of the reference object: That is, *in front of the house* refers to the intrinsic front orientation of the house. Due to (43), *John is sitting in front of the house* is true just in case 'John is sitting and John's region intersects the front orientation of the house.'[3]

3. Symmetric Locatives

Symmetric locatives include PPs with *across*, *through*, *over*, *around*, and *past*, and their semantics crucially involve *betweenness* relation. In section 1, we noted some semantic facts revealing the nature of symmetric locatives. This section proposes lexical semantics of *across* and *through*, then characterizes semantically the class of symmetric locatives. Now (44) and (45) are instantiations of Semantic Rule–1, and define the meaning of *across* and *through* as a unary preposition.

(44) *across* α:
For *m* a one-place motion predicate, and for α a noun phrase denoting an individual, interpret the one-place predicate *m+across+*α, as follows:
(**across**(α))(**m**)(x) = 1 iff **m**(x) and TRAV(x, π, T),
where Between(®(α), π_s, π_g), and for some $t\in$T,
ON(®(x,t), ®(α,t))

[3] Locatives may refer to an orientation which is not an intrinsic one of the reference object. For example, *John was sitting across the street* contains a symmetric locative which do not refer to an inherent orientation of the street, but the PP requires a locative perspective (point of view) to determine a relevant orientation. See Nam (1995) for a formal treatment of deictic locative perspective.

(45) *through*: α:
For *m* a one-place motion predicate, and for α a noun phrase denoting an individual, interpret the one-place predicate *m+through+*α, as follows:
(**through**(α))(**m**)(x) = 1 iff **m**(x) and TRAV(x, π, T), where Between(®(α), π_s, π_g), and for some t∈T, IN(®(x,t), ®(α,t))

The only difference between *across* and *through* defined above is their last condition on the intermediate location of the moving object: ON(®(x,t), ®(α,t)) for *across* α and IN(®(x,t), ®(α,t)) for *through* α. Thus, *across* requires the object x to be "on" the reference object, whereas *through* requires x to be "in" the reference object. The following examples illustrate how those PPs are interpreted in a sentence, and we see that (46) entails 'John was on the street,' and (47) 'John was in the tunnel.'

(46) *John ran across the street* is true iff
for some interval T<**now**, **run** (**john**)(T) and TRAV(**john**, π, T)
for a path π where Between(®(**the street**), π_s, π_g) and
for some t∈T, ON(®(**john**, t), ®(**the street**, t)).

(47) *John drove through the tunnel* is true iff
for some interval T<**now**, **drive**(**john**)(T) and TRAV(**john**, π, T)
for a path π such that Between(®(**the tunnel**), π_s, π_g), and
for some t∈T, IN(®(**john**, t), ®(**the tunnel**, t)).

In the previous section, two general interpretative rules are given for different types of locatives combining with motion verbs and stative verbs. Now we account for the symmetry effects illustrated in 1.1-1.4 by identifying some unique characteristics of the semantics of symmetric locatives. We characterize symmetric locatives as ones determining a set of paths which is closed under "path-converse" relation. That is, if a path π is determined by a symmetric locative *f*, then π^{-1} is also determined by *f*.

(48) Definition: *Symmetric Locatives*
For *f* a locative modifier, *f* is a *symmetric locative* iff
f determines a set $\prod_f$ of paths such that $\forall\pi\in\prod_f$, $\pi^{-1}\in\prod_f$,
i.e., $\prod_f$ is closed with respect to "path-converse" relation.

For example, as we see in (46), the PP *across the street* determines a path π such that Between(®(**the street**), π_s, π_g), and since Between is symmetric on second and third arguments, Between (®(**the street**), π_g, π_s). Then by definition, $\pi_s = \pi^{-1}{}_g$, and $\pi_g = \pi^{-1}{}_s$, and so Between (®(**the street**), $\pi^{-1}{}_s$, $\pi^{-1}{}_g$). Therefore, the same PP also determines π^{-1}, thus symmetric. This definition now identifies the class of locatives which induces the symmetric effects illustrated in 1.1-1.4.

4. *Directional* Locatives and Homogeneous Paths

Prepositional phrases with *to, from, into, out of,* and *towards* are called *directional locatives*, which indicate directions of movement by referring to the source/goal of a path. Thus, *to*, *into*, and *towards* refer to the goal of a path, but *from* and *out of* refer to the source of a path. *To* and *from* are semantically more basic than *into* and *out of* in the sense that the latter not only specifies the goal/source of a path but indicate the goal/source are in

the interior region of a reference object. The following entailment patterns illustrate the difference.

(49) a. *Cindi walked to/from the market* does not entail
b. ⊭ *Cindi was in the market*

(50) a. *Cindi walked into/out of the market* entails
b. ⊨ *Cindi was in the market*

The contrast supports the intuition that, unlike *out of, from* takes a reference object as a point abstracted from a volume or an area. The following contrast between *from* and *out of* in (51-54) also suggests that *out of* takes a reference object as a container with a non-trivial interior region.

(51) a. *The prisoner escaped from his guards!*
b. **The prisoner escaped out of his guards!*

(52) a. *His friend saved him from a lion*
b. **His friend saved him out of a lion*

(53) a. *He guards his master from enemies*
b. **He guards his master out of enemies*

(54) a. *I put the book out of the box*
b. **I put the book from the box*

As we mentioned in 1.5, *to* and *towards* induce a difference in aspectual interpretation: *telic* vs. *atelic*. Thus, in Hinrichs (1986), (55a) and (55b) are classified into Vendler's classes of *accomplishment* (telic) and *activity* (atelic), respectively.

(55) a. *Fangs slithered to the rock*
b. *Fangs slithered toward the rock*

Now let us define the lexical semantics of *to* and *towards*, and characterize their semantic difference in terms of homogeneity of paths. A locative PP with *to*, as defined in (56), determines a set of paths which share a specific goal region denoted by the reference object. Thus, due to the semantic rule–1, we interpret the PP combining with a motion predicate as in (57), and this interprets the sentence *John ran to the post office* as in (58). The truth conditions in (58) guarantees that the sentence entails (59).

(56) *to* α: determines a set of paths π such that
π_g = ®(α) and $\pi_g \cap \pi_s = \emptyset$.

(57) Let *m* be a one-place motion verb, and α an NP denoting an individual, interpret the VP *m+to+*α, the concatenation of *m* and the PP*to* α, by:
(**to** α)(**m**)(x) iff **m**(x) and TRAV(x,π,T)
where π_g = ®(α) and $\pi_g \cap \pi_s = \emptyset$.

(58) *John ran to the post office* is true iff
there is an interval T<**now**, and a path π such that **ran** (**john**, T) and TRAV(**john**, π, T) where π_g = ®(**the post office**) and $\pi_g \cap \pi_s = \emptyset$.

(59) *John was at the post office*

Directional locatives with *towards* get interpreted quite differently from those with *to*. First of all, a PP with *towards*, like *towards the post office*, does not refer to the goal of the movement, but simply specifies the direction of the movement: That is, the movement is directed toward the post office. In other words, the distance from the moving object to the post office is continuously decreasing. We interpret this then in terms of "relative nearness" relation. Thus, *towards* α is interpreted as determining a path defined as in (60).

(60) *towards* α: determines a path π such that for all $i \leq j \in$Domain(π), Nearer(π(j), ®(α), $\pi(i)$).

(61) *John ran towards the post office* is true iff
for an interval T<**now**, and for a path π,
ran(**john**, T) and TRAV(**john**, π, T), where
for all $i \leq j \in$Domain(π), Nearer($\pi(j)$, ®(**the post office**), $\pi(i)$).

Then, Semantic Rule–1 interprets the sentence in (61) as given, and we can see that the sentence does not entail (59) *John was at the post office.*

Now the difference between *to* in (56) and *towards* in (60) can be identified in terms of "homogeneous" paths defined below:

(62) Definition: *Homogeneous Paths*
For *f* a locative PP, *f determines a homogeneous set* $\prod_f$ *of paths*
iff $\forall\pi \in \prod_f$, if π' is a subpath of π, $\pi' \in \prod_f$,
i.e., $\prod_f$ is closed with respect to "subpath" relation.

From the definition (62), we see that the PP *towards the post office* determines homogeneous paths, but *to the post office* does not. As we mentioned earlier, Hinrichs (1985) adopts Vendler's (1967) temporal criteria (*continuity* vs. *punctuality* and *homogeneity* vs. *heterogeneity*) for his four verb classes, and he characterizes the difference between *to* and *towards* in terms of *temporal heterogeneity/homogeneity*: Hinrichs gives a lexical account of the aspectual difference with the notion of spatio-temporal location and Carlson's (1977) dichotomy of stage level and individual level predicates. Hinrichs paraphrases Vendler's (1967:101) ideas as follows:

> "Since doing something *for* x amount of time means doing something during most if not all subintervals of the interval x, sentences such as (61), which refers to atelic events or activities, can be characterized as being *temporally homogeneous*....
>
> To do something *in* x amount of time, on the other hand, means to do something at some unique interval within x. Since telic events or accomplishments can be modified by temporal *in*, they, in contrast to activities or atelic events, can be described as being temporally heterogeneous, ..." (Hinrichs 1986:349)

Vendler characterizes the difference in terms of *temporal heterogeneity* vs. *temporal homogeneity*, but here we associate the aspectual difference with the difference among paths that the locatives determine. The parallel definition given in (62) reveals an intuitive correspondence between the temporal homogeneity of events and the spatial homogeneity of paths.

We can show that, due to the homogeneity of the paths determined by *towards the post office*, the interpretation in (61) above implies the temporal homogeneity of the event: That is, if the sentence is true for some past interval T<**now**, then for most if not all subintervals T' of T, the sentence is true. But we can notice quickly that PPs with *to* do not determine homogeneous paths, and so the interpretation of (58) *John ran to the post office* does not imply temporal homogeneity of the event.

References

Bennett, D. (1975) *Spatial and Temporal Uses of Prepositions: An Essay on Stratificational Semantics*. London. Longman

Benthem, J. v. (1983) *The Logic of Time*. 1/2nd edition, Kluwer Academy Publishers

Carlson, G. (1977) *Reference to Kinds in English*, Doctoral dissertation. University of Massachusetts, Amherst

Cresswell, M. (1978) 'Prepositions and Points of View,' *Linguistics and Philosophy* 2: 1-41

Cresswell, M. (1978b) 'Adverbs of Space and Time,' in Guenthner, F. and S.J. Schmidt (eds.) *Formal Semantics and Pragmatics for Natural Language*, Reidel, Dordrecht.

Herskovits, A. (1986) *Language and Spatial Cognition*. Cambridge University Press

Hinrichs, E. W. (1985) *A Compositional Semantics for Aktionsarten and NP Reference in English*, Doctoral dissertation. The Ohio State University

Hinrichs, E. W. (1986) 'A Compositional Semantics for Directional Modifiers–Locative Case Reopened,' in *COLING* 86: 347-349, Association for Computational Linguistics, Bonn

Jackendoff, R. (1983) *Semantics and Cognition*. The MIT Press

Jackendoff, R. (1990) *Semantic Structures*. The MIT Press

Kamp, H. (1979) 'Events, Instants, and Temporal References,' in R. Bäuerle and U. Egli and A. von Stechow (eds.) (1979) *Semantics from Different Points of View*, Springer-Verlag, Berlin

Keenan, E.L. & Faltz, L. M. (1985) *Boolean Semantics for Natural Language*. D. Reidel, Dordrecht

Miller, G. & Johnson-Laird, P. N. (1976) *Language and Perception*. Cambridge, Harvard U. Press

Montague, R. (1970) 'English as a Formal Language,' in Bruno Visentini *et al. Linguaggi nella Societa e nella Tecnica*. Milan, Edizioni di Comunita. Reprinted in R. Thomason (ed.)(1974) *Formal Philosophy*, Yale University Press, New Haven

Montague, R. (1973) 'The Proper Treatment of Quantification in Ordinary English,' in Hintikka, J. Moravcsik, & P. Suppes (eds.) *Approaches in Natural Language*, Reidel, Dordrecht

Nam, S. (1995a) *The Semantics of Locative Prepositional Phrases in English*. Doctoral dissertation, University of California, Los Angeles.

Nam, S. (1995b) 'Argument Orientations of Locative PPs in English,' in J. Fuller, H. Han, and D. Parkinson (eds.) *Proceedings of the Eleventh Eastern States Conference on Linguistics*. Ithaca, NY: DMLL Publications

Parsons, T. (1990) *Events in the Semantics of English: A Study in Subatomic Semantics*. The MIT Press, Cambridge
Piaget, J. and Inhelder, B. (1948/1968) *The Child's Conception of Space*. The Norton Library, New York
Rescher, N. & Garson, J. (1968) 'Topological Logic,' The Journal of Symbolic Logic vol. 33.4: 537-548
Robinson, R. M. (1959) 'Binary Relations as Primitive Notions in Elementary Geometry,' in L, Henkin et al (eds.)
Sondheimer, N. (1978) 'A Semantic Analysis of Reference to Spatial Properties,' *Linguistics and Philosophy* 2: 235-280
Talmy, L. (1983) 'How Language Structures Space,' in H. Pick & L. Acredolo (eds.) *Spatial Orientation: Theory, Research, and Application*. New York, Plenum
Tarski, A. (1927/1956) 'Foundations of the Geometry of Solids,' in Tarski (1956) *Logic, Semantics, Mathematics*. Oxford U. Press. translated from French (1927)
Tarski, A. (1959) 'What is Elementary Geometry?' in Henkin, L. et al (eds.)(1959)
Vendler, Z. (1967) *Linguistics in Philosophy*. Ithaca, Cornell University Press

Aspectual Composition and Dynamic Logic[1]

Ralf Naumann
University of Düsseldorf

(1) The Interpretation of Verbs

Event-based approaches like the one of Krifka as well as the index-based approach of Verkuyl are characterized by the fact that there are only minimal differences between the interpretations of verbs in the lexicon. In Krifka's approach all non-stative verbs are basically predicates of events, P(e), that are characterized by having cumulative reference. If the predicate P applies to an event e and also to an event e', the sum (join) e ⊔ e' of the two events satisfies the predicate as well. According to Verkuyl, the class of non-stative verbs is characterized by having the feature [+ ADD TO]. At the formal level this feature is represented by an index-argument i plus some further conditions on i that have to do with the way in which elements of the set denoted by an argument NP are correlated with the indice structure generated by the verb.

A consequence that is common to both approaches is that the notion of an aspectual class, as it has been used e.g. by Vendler, is not applicable to the verbal level but only, if at all, to higher levels, for instance, the VP-level. This is in contrast to the approach of Dowty who applies the notion of aspectual class already to the verbal level.

From what has been said so far it follows that there is no distinction between non-stative verbs. Differences at the VP-level must consequently be attributed to differences with respect to the semantic properties of the argument NPs. Suppose that from an aspectual point of view there are basically two types of NPs. First, there are NPs that have cumulative (homogeneous) reference like bare plurals (*apples*) and mass NPs (*water*, *ice*), whereas the second class is characterized by having quantized (heterogeneous) reference. To this class belong NPs like *five apples*, *at most five apples* and *more than five apples*. Consequently one should expect that at the VP-level an expression can be distinguished by the fact of whether it has an internal NP with cumulative or with quantized reference. Consider the following examples.

(1)	a.	John ate an apple	=>	terminative
	b.	John crossed the street	=>	terminative
	c.	John pushed the cart	=>	durative
	d.	John walked	=>	durative
	e.	John pushed the cart to the station	=>	terminative
	f.	John walked three miles	=>	terminative
	g.	John hit Mary	=>	terminative

A sentence is said to have a terminative interpretation if it is possible to modify it with an *in*-adverbial, whereas the interpretation is durative if the sentence can be modified by a *for*-adverbial.

The NPs occurring in the sentences can all be said to be [+ SQA] in the sense of

[1] I would like to thank Yoshiki Mori for his helpful comments and suggestions.

Verkuyl or to have quantized reference in the sense of Krifka. Consequently any difference in aspectual behaviour cannot be attributed to a difference in the semantic properties of these NPs. For Krifka, the difference between (1a,b) and (1c,d) can be explained as a difference with respect to the thematic relations that are assigned to the argument-positions of the verbs. These relations have different properties that are aspectually relevant. In the lexicon a (transitive) verb is not simply a predicate of events, $\lambda eP(e)$, but rather has the form given in (2). (θ_1 and θ_2 are thematic relations like Agent or Patient.)

(2) $\lambda y \lambda x \lambda e[P(e) \wedge \theta_1(e,x) \wedge \theta_2(e,y)]$[2]

The difference between verbs like *eat* and *cross*, on the one hand, and verbs like *push* and *walk* on the other can then be explained in the following way. The properties of the different thematic relations induce an asymmetry between the different argument positions: there is a distinction between those arguments that are incremental and those that are not.

- eat(y)(x), cross(y)(x) => y is incremental
- push(y)(x), walk(x) => neither x nor y are incremental

Incrementality is defined as graduality of the thematic role which in turn is defined by two mappings between the object and the event-domain and uniqueness of objects (see Krifka(92)) for details). If an argument position is incremental, there will be a transfer of referential properties from the argument NP to the whole expression (mostly the VP). Note that the addition of thematic relations to the representation of a verb does not change its referential properties. An expression of the form given in (2) is still cumulative because all thematic relations are assumed to be summative such that if the predicate applies to an event e and objects x,y as well as to an event e' and objects x',y', then θ_1 and θ_2 apply to the joins $x \sqcup x'$ and $y \sqcup y'$ of x, x' and y, y', respectively, as well. The reason why the difference with respect to the properties of thematic relations is not taken as an aspectual difference is that only the referential properties of the whole expression are taken into consideration and in this respect there is no difference between verbs like *eat* and *push*, as was just shown. Thus it seems that Krifka can account for the difference in aspectual behaviour of the sentences in (1a,b) as opposed to the sentences in (c,d). Yet there remain at least the following three problems.

(i) In an intuitive sense verbs like *push* and *walk* are gradual in a way similar to *eat* and *cross* and are opposed to verbs like *see* or *love*. Empirically this difference is manifested in the different behaviour with respect to the progressive. Whereas *John is walking* or *John is pushing the cart* are perfectly acceptable, it is not in general possible to use the progressive with *see* or *love*: **John is seeing Mary* and **John is loving Mary*.

(ii) According to the analysis given above, graduality of one thematic role is a necessary condition for a VP to have quantized reference (i.e., to be terminative). But the example in (1g) shows that this is not correct. A verb like *hit* is not gradual, yet sentence (1g) is terminative. (Below it will be shown that the notion of graduality is even inadequate to account for the aspectual behaviour of verbs like *eat*; see section 4.)

[2] Krifka does not bind the two nominal arguments x and y by λ-abstraction but treats them as free variables. This leads to problems, as pointed out by Verkuyl; see Verkuyl(93), section 11.4.

(iii) Granting that verbs like *push* and *walk* are distinguished from verbs like *eat* by having only constant roles, there is the problem of how these verbs can be distinguished from stative verbs like *love* or *be ill*.

Let us first tackle problem (i). In what sense exactly are verbs like *push* and *walk* gradual as opposed to, say, *see*? Consider the following two sentences given in (3).

(3) a. John walked for two minutes.
b. John saw Mary for two minutes.

In the case of (3a) there is something which constantly changes during the period of walking, namely, the location (or position) of the subject, i.e., John. The same does not hold if John sees Mary for the same period. Once Mary is seen, which can be taken to represent a change from John's not being aware of Mary to his being aware of her, there is no further change with respect to his being aware of this object. (Of course, John can become aware of the colour of Mary's eyes and of many other things. But these represent different kinds of changes.) For walking, on the other hand, the kind of change that is brought about at the beginning still holds when John continues walking.

The above considerations can be interpreted in such a way that non-stative verbs express changes with respect to one of their arguments.

- eat(y)(x) => change with respect to y
- cross(y)(x) => change with respect to x
- push(y)(x) => change with respect to y
- walk(x) => change with respect to x

What exactly is changed? The quantity with respect to which a change effected by the verb occurs is in most cases not explicit, but there are paraphrases that give an indication as to what undergoes a change. Consider the following pairs of sentences.

(4) a. The colour of John's hair changed.
b. John dyed his hair.

(5) a. Der Preis des Goldes stieg um einen Dollar. (The price of gold rose by one dollar)
b. Gold verteuerte sich. (Gold became more expensive)

(6) a. My home used to be in Maryland but now it is in Los Angeles.
b. Barbara moved from Maryland to Los Angeles.

(7) a. The colour of the wall is now blue.
b. John painted the wall blue.

In (4b) it is the colour of the hair that undergoes a change, as shown by the paraphrase in (4a). The colour can be taken to be a property of the hair. In (5b) it is the price of the gold that changed, where again the former can be interpreted as one of the properties of the latter. A similar argument can be adduced in the two remaining cases. According to these data, one can claim that it is a property of one object denoted by the verb that undergoes a change in the sense that at each moment this property has a particular value and that for different moments the values can be different. A verb denotes then a change of this value. For example, the verb *dye* denotes a change of value with respect to colour, whereas *walk* denotes a change of value with respect to the location (or position) of an object.

The notion of a property must be understood in a broad sense, including also what may be called aspects or roles of objects. Consider the following three examples.

(8) a. John read a book.
b. Mary peeled the apple.
c. John was elected president.

It can be argued that in all three cases no property of an object is directly involved. Only in the second case seems it possible to argue that it is the property of having a surface that is involved, but in the other two cases no such argument is possible. Krifka, in discussing the first two examples, calls that which is involved an aspect of the object, the pages of the book in the first case and the surface of the apple in the second. (Krifka(92:45)) What gets changed in the case of (8b) is the particular surface the apple has. In the third example it is a certain role or function of John that is at stake. Here John acquires the role (function) of being president. This example also gives an indication of how the first example can be analysed. An aspect of an object is a particular part (or are particular parts) of it that get a role by being involved in a process denoted by a verb. In the case of reading a book the parts are the pages of the book. The role they have cannot be described independently of the verb, i.e. *read*, itself: it is just the role of being read.

Because of examples like those in (8) I will not only speak of properties but also of roles and aspects. The thesis then is that (non-stative) verbs denote changes with respect to a property (role or aspect) of one object that is an argument of the verb.

- eat(y)(x) => change with respect to y or the mass of y
- cross(y)(x) => change with respect to the location of x
- push(y)(x) => change with respect to the location of y
- walk(x) => change with respect to the location of x

The difference between verbs like *eat* or *cross* and verbs like *push* and *walk* must then be explained as a difference with respect to the kind of change denoted. Note that on this view too there is an aspectual asymmetry: there is a distinction between the argument with respect to which a change occurs and those arguments for which this does not hold. Yet there are the following two differences between the perspective adopted here and the one used by Krifka. The first concerns the argument that is distinguished by this asymmetry. In the case of *cross*, e.g., it is the internal argument that is incremental on Krifka's view, whereas it is the external argument on the present perspective. More important is the second difference. Recall that a verb like *push* does not show any aspectual asymmetry because both arguments are assigned a constant role. This leads to the third problem above of how these verbs can be distinguished from stative verbs. On the present view there is an asymmetry with respect to the two arguments of *push* because it is the internal argument with respect to which a change matters. The case is different for stative verbs. In the case of *John is ill* or *Mary loves John* no change is involved at all, i.e., neither with respect to John nor with respect to Mary. Furthermore distinguishing stative verbs from non-stative ones by means of the notion of change accounts for the intuition discussed above that verbs like *walk* are gradual in a way similar to verbs like *eat* or *cross*.

(2) The Interpretation of Accomplishment-Verbs

In computer science changes are modeled by programs which are interpreted as relations between states. In dynamic semantics this idea is taken up by interpreting verbs and sentences as relations between (information) states or as generalized quantifiers of states (as in DPL and DMG). Yet verbs are uniformly interpreted as tests, that is, as programs where the input-state is identical to the output state. This makes verbs essentially static because actually there is no change involved.

It is obvious that if verbs denote 'real' changes in the sense described above, they can no longer be interpreted as tests. But what kind of program shall an accomplishment-verb like *eat* denote? One will come closer to a solution if one considers one way of describing what it means to eat an apple. Suppose the apple is cut into n pieces. Then the eating of the apple can be described in the following way.

(9) While the number of pieces is not zero continue eating.

The description in (9) can be interpreted as a kind of instruction or command. It says how one has to proceed in order to eat an apple. On the other hand, this description has exactly the structure of a **while**-loop, a program-construct used in imperative programming-languages.

(10) **while** ϕ **do** π

This (complex) program consists of three different components.

(i) a boolean condition ϕ
(ii) a program π
(iii) a control structure **while do**

The intuitive meaning of this program is that the subprogram π is executed as long as the condition ϕ holds. After each execution of π it is tested whether ϕ still evaluates to true. In case it does, π is executed another time, otherwise the whole program terminates. The boolean condition ϕ can therefore be interpreted as specifying a terminating or stopping condition in the sense that as soon as $\neg\phi$ holds, the loop terminates. The real dynamic component is the program π because it brings about a change with respect to the memory of the computer.

Interpreting an accomplishment-verb like *eat* as denoting a **while**-loop is possible only if a decompositional analysis is chosen because each of the three components must be modeled as part of the interpretation of such a verb.

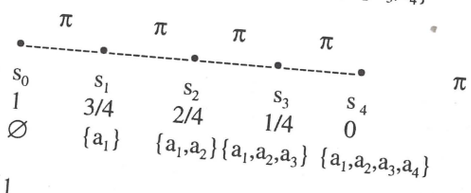
The first question that must be answered is how the program π should be modeled. In section (1) it was argued that what is changed is the value of a property of one object denoted by an argument of the verb. Such a property is therefore a kind of dynamic property the value of which can change due to external influences.

From a programming point of view such a dynamic property can be represented by a store (program variable, pigeon-hole). In each state of a computation this store has a particular value. This value can be changed by an assignment command. The effect of an execution of such a command is that the store has different values in the input-state and in the output-state. Applied to the case of a verb this means that the property represented by a

store has a particular value in the input-state that is changed to a different value in the output-state. Applying this consideration to the eating of an apple that is cut into n pieces this means that the assignment command will model the transition from a situation in which already k pieces have been eaten to a situation where k+1 pieces are eaten. The property that is affected is the mass of the apple because each piece uniquely corresponds to a part of the mass of the apple. The store, say u, can therefore be taken to represent the mass of the apple. In the input state of an execution of the assignment command the store u has as its value the set of pieces that have already been eaten and in the output-state the value is the union of this set and the singleton containing the piece just eaten.

In the initial state of a computation the value of u will be the empty set corresponding to the fact that in this state no piece has been eaten. Then the n pieces are successively eaten. The computation terminates when the last piece has been eaten, that is, if there are no more pieces.[3] In terms of the mass of the apple an alternative description is the following. The value of u in the initial state is the total mass of the apple, which can be taken to correspond to the value 1. Each execution of the assignment command reduces this mass about the mass of the piece that is eaten. The computation stops if the mass is zero, i.e., if the apple has vanished. As an illustration a situation will be taken in which the apple is cut into four pieces a_1, a_2, a_3 and a_4.

states	mass	pieces eaten
s_0	1	$\varnothing$
s_1	3/4	$\{a_1\}$
s_2	2/4	$\{a_1,a_2\}$
s_3	1/4	$\{a_1,a_2,a_3\}$
s_4	0	$\{a_1,a_2,a_3,a_4\}$



$\pi \equiv u := u \oplus 1$

Figure 1

In this particular situation five states are involved, i.e., one gets the execution sequence $\sigma=s_0...s_4$. The values of u are linearly ordered. For instance, if the representation in terms of the apples eaten is chosen, one has $s_i(u) \subseteq s_{i+1}(u)$. The assignment command can then be modeled as set union: $s_{i+1}(u) = s_i(u) \cup \{a_i\}$.

As the discussion so far has shown, there are different ways in which the eating of an apple can be described by using a **while**-loop. Which of the representations should be chosen? In order to answer this question one has to take into account that the representation must apply not only to a verb like *eat* but also to other accomplishment-verbs like *cross* or *read*. In the case of *cross*, say, *cross the street*, the property that is at stake is the location of the subject. This property can be modeled by another store, say, u'. The possible values of u' can be taken to be paths in the following sense. In the initial state John's location will be a point on one side of the street, say, l_0. So far, he hasn't done a single step such that the path traversed so far is the empty path. If John starts moving, his location is constantly changing (disregarding rests) and consequently the path traversed is getting longer. Now the path John has traversed if he is at l_n can be represented by the set of all locations he has already passed, starting from l_0. Thus the path traversed in l_n will be the set $\{l_0, ..., l_n\}$. Another way of looking at the relation between the locations passed by John and the path traversed by him is to identify the location l_n with the path the initial point of which is l_{n-1} and whose end point is l_n. The assignment command can then be interpreted as path concatenation. The value of store u' in the input-state is the path already traversed whereas the value of u' in the output-state is the concatenation of this path with the path that corresponds to the next location that is reached by John. What this next location is depends on the measure that is chosen. For instance the crossing can be measured in single steps, several steps, in yards or in meters. The choice of a measure is context-dependent. This dependency is exactly similar to the dependency in the case of eating where the partition of the apple depends on the particular context.

Thus the values of the active variable can be of different sorts (or types): pieces (of objects), numbers or paths, for example. Depending on the sort of object, there is a corresponding difference with respect to the particular operation that is needed to compute the value of the variable in the output-state. What is needed therefore is a generalization which is based on what is common to all the different assignment commands. One arrives at such a generalization by the following consideration. The nature of the particular values assigned to the different stores does not matter with respect to the dynamic structure generated by an execution of the program. What is important is that each execution is done with respect to a particular set that is linearly ordered (i.e., a chain). What these elements are is unimportant. It is therefore possible to identify them with natural numbers. It follows that each particular ordering for a particular verb can be embedded in the set of natural numbers. Take as an example the apple that is cut into n pieces. In this case the values wqill correspond to the first n natural numbers.

Interpreting the values of the active variable to be natural numbers has the further advantage of allowing the operation of how the value in the output-state is computed to simply be addition: $u := u \oplus 1$. Each computation starts in a state where the (active) store u has value 0 and by each execution of the command, i.e., each execution of the loop, the value is raised by 1, until finally the terminating condition is reached for a particular n. In the case of eating an apple, e.g., the value n corresponds to the value of the mass of the apple being 0, whereas in the case of reading a book the correspondence can be taken to be between the n pages of the book and the n first natural numbers.

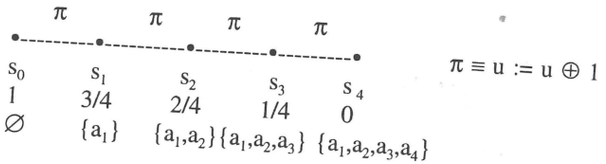
The assignment command that is used is underdetermined in at least two respects. First, the exact interpretation of the $\oplus$-operation is not determined and second, the exact 'nature' of the values is left open. Furthermore, the particular chain with respect to which a computation is done is given by the context (environment). The linear ordering underlying a computation is similar to the notion of a path as it is used in localistic theories and the approach of Verkuyl.

[3] One may object that it is possible to eat two pieces at the same time and that this cannot be accounted for by the assignment command. This is not correct. Each computation is done with respect to a particular subdivision of the apple. Each of these subdivisions can be taken to form a linear sequence in the sense that to each execution of the assignment command there corresponds exactly one part of the subdivision. A case as described in the objection can always be linearized by interpretating it as a computation with respect to a subdivision in which the two pieces form a single piece. If one interprets the subdivision of the apple as a partition, this linearization amounts to treating each cell of the original partition as one piece.

store has a particular value in the input-state that is changed to a different value in the output-state. Applying this consideration to the eating of an apple that is cut into n pieces this means that the assignment command will model the transition from a situation in which already k pieces have been eaten to a situation where k+1 pieces are eaten. The property that is affected is the mass of the apple because each piece uniquely corresponds to a part of the mass of the apple. The store, say u, can therefore be taken to represent the mass of the apple. In the input state of an execution of the assignment command the store u has as its value the set of pieces that have already been eaten and in the output-state the value is the union of this set and the singleton containing the piece just eaten.

In the initial state of a computation the value of u will be the empty set corresponding to the fact that in this state no piece has been eaten. Then the n pieces are successively eaten. The computation terminates when the last piece has been eaten, that is, if there are no more pieces.[3] In terms of the mass of the apple an alternative description is the following. The value of u in the initial state is the total mass of the apple, which can be taken to correspond to the value 1. Each execution of the assignment command reduces this mass about the mass of the piece that is eaten. The computation stops if the mass is zero, i.e., if the apple has vanished. As an illustration a situation will be taken in which the apple is cut into four pieces a_1, a_2, a_3 and a_4.

states	mass	pieces eaten
s_0	1	$\varnothing$
s_1	3/4	$\{a_1\}$
s_2	2/4	$\{a_1,a_2\}$
s_3	1/4	$\{a_1,a_2,a_3\}$
s_4	0	$\{a_1,a_2,a_3,a_4\}$



$\pi \equiv u := u \oplus 1$

Figure 1

In this particular situation five states are involved, i.e., one gets the execution sequence $\sigma=s_0...s_4$. The values of u are linearly ordered. For instance, if the representation in terms of the apples eaten is chosen, one has $s_i(u) \subseteq s_{i+1}(u)$. The assignment command can then be modeled as set union: $s_{i+1}(u) = s_i(u) \cup \{a_i\}$.

As the discussion so far has shown, there are different ways in which the eating of an

[3] One may object that it is possible to eat two pieces at the same time and that this cannot be accounted for by the assignment command. This is not correct. Each computation is done with respect to a particular subdivision of the apple. Each of these subdivisions can be taken to form a linear sequence in the sense that to each execution of the assignment command there corresponds exactly one part of the subdivision. A case as described in the objection can always be linearized by interpretating it as a computation with respect to a subdivision in which the two pieces form a single piece. If one interprets the subdivision of the apple as a partition, this linearization amounts to treating each cell of the original partition as one piece.

apple can be described by using a **while**-loop. Which of the representations should be chosen? In order to answer this question one has to take into account that the representation must apply not only to a verb like *eat* but also to other accomplishment-verbs like *cross* or *read*. In the case of *cross*, say, *cross the street*, the property that is at stake is the location of the subject. This property can be modeled by another store, say, u'. The possible values of u' can be taken to be paths in the following sense. In the initial state John's location will be a point on one side of the street, say, l_0. So far, he hasn't done a single step such that the path traversed so far is the empty path. If John starts moving, his location is constantly changing (disregarding rests) and consequently the path traversed is getting longer. Now the path John has traversed if he is at l_n can be represented by the set of all locations he has already passed, starting from l_0. Thus the path traversed in l_n will be the set $\{l_0, ..., l_n\}$. Another way of looking at the relation between the locations passed by John and the path traversed by him is to identify the location l_n with the path the initial point of which is l_{n-1} and whose end point is l_n. The assignment command can then be interpreted as path concatenation. The value of store u' in the input-state is the path already traversed whereas the value of u' in the output-state is the concatenation of this path with the path that corresponds to the next location that is reached by John. What this next location is depends on the measure that is chosen. For instance the crossing can be measured in single steps, several steps, in yards or in meters. The choice of a measure is context-dependent. This dependency is exactly similar to the dependency in the case of eating where the partition of the apple depends on the particular context.

Thus the values of the active variable can be of different sorts (or types): pieces (of objects), numbers or paths, for example. Depending on the sort of object, there is a corresponding difference with respect to the particular operation that is needed to compute the value of the variable in the output-state. What is needed therefore is a generalization which is based on what is common to all the different assignment commands. One arrives at such a generalization by the following consideration. The nature of the particular values assigned to the different stores does not matter with respect to the dynamic structure generated by an execution of the program. What is important is that each execution is done with respect to a particular set that is linearly ordered (i.e., a chain). What these elements are is unimportant. It is therefore possible to identify them with natural numbers. It follows that each particular ordering for a particular verb can be embedded in the set of natural numbers. Take as an example the apple that is cut into n pieces. In this case the values wqill correspond to the first n natural numbers.

Interpreting the values of the active variable to be natural numbers has the further advantage of allowing the operation of how the value in the output-state is computed to simply be addition: $u := u \oplus 1$. Each computation starts in a state where the (active) store u has value 0 and by each execution of the command, i.e., each execution of the loop, the value is raised by 1, until finally the terminating condition is reached for a particular n. In the case of eating an apple, e.g., the value n corresponds to the value of the mass of the apple being 0, whereas in the case of reading a book the correspondence can be taken to be between the n pages of the book and the n first natural numbers.

The assignment command that is used is underdetermined in at least two respects. First, the exact interpretation of the $\oplus$-operation is not determined and second, the exact 'nature' of the values is left open. Furthermore, the particular chain with respect to which a computation is done is given by the context (environment). The linear ordering underlying a computation is similar to the notion of a path as it is used in localistic theories and the approach of Verkuyl.

As was said above, the function of the boolean condition ϕ is to determine a terminating condition. From this it follows that ϕ must involve the active variable because it is only then that there is relation between ϕ and the assignment command. In general ϕ will have the form $u \neq \mathbf{b}$, where **b** is a specific value that must hold for u in the final state of a computation. For instance, in the case of *eat* the value will be 0 corresponding to the fact that the object eaten has vanished which can be represented by the value of its mass having become 0. For *cross* the value of u' determined by ϕ is a particular point on one side of the object that is crossed, e.g., a particular point on the other side of the street (for a particular computation).

The implementation will be done in Muskens' 'Logic of Change' (LoC). (For details on LoC, see Muskens(94) and Naumann(95)) LoC is a three sorted Type Logic Ty3. Besides the usual types of truth values t and ordinary objects of type e, there are the type s of states and the type π of pigeon-holes (stores, program variables). A state, i.e. an object of type s, can be thought of as a list that consists of the (current) values of all stores. The relation between a state, the set of stores and the values of the stores in this state is expressed by a (non-logical) constant V of type $(\pi,(s,e))$ that assigns to each store and state the value of this store in that state.

In QDL the **while**-loop is simulated by means of sequencing ;, the test command ? and iteration (Kleene-star) *: $((\neg\phi)? ; \pi)^* ; \phi?$. The first two can be translated into Type Theory as given in (11).

(11) a. $\pi ; \pi' = \lambda ij\exists k[\pi^\circ(i)(k) \wedge \pi'^\circ(k)(j)]$[4]
b. $\phi? = \lambda ij[i=j \wedge \phi^\circ(j)]$

(here π°,π'° and ϕ° are the translations of π,π' and ϕ in Type Theory, respectively)
The iteration-command is not first-order definable. One either needs higher-order quantification or infinite disjunction. Extending Type Theory by the latter device allows to translate iteration as given in (12).

(12) $\pi^* = \bigvee_{n\in\omega} \pi^n$
$= \lambda i\lambda j[i=j \vee \pi^\circ(i)(j) \vee \exists k[\pi^\circ(i)(k) \wedge \pi^\circ(k)(j)] \vee \exists kl[\pi^\circ(i)(k) \wedge \pi^\circ(k)(l) \wedge \pi^\circ(l)(j)] \vee ...]$

(12) will be abbreviated to (13a) which will be further abbreviated to (13b).

(13) a. $= \lambda ij\exists n\exists s_0...s_n[i=s_0 \wedge j=s_n \wedge \forall k{:}0\leq k<n\ \pi^\circ(s_k)(s_{k+1})]$
b. $= \lambda ij\exists n\exists s_0...s_n[i=s_0 \wedge j=s_n \wedge s_0\pi^\circ{*}s_n]$

The translation of the assignment command u := u ⊕ 1 is given in (14a). Application to states i and j will be abbreviated as in (14b).

(14) a. $[u := u \oplus 1] = \lambda ij[V(u)(j) = V(u)(i) \oplus 1 \wedge \forall v[v\neq u \rightarrow V(v)(i) = V(v)(j)]]$
b. $i[u := u \oplus 1]j$

[4] In the sequel the following convention is used: $i,k.,l,s_i$ ($i\in\mathbb{N}$) are variables of type π, u,u' are constants of type π.

The interpretation of *eat* is given in (15).

(15) eat ⇝
$\lambda v'\lambda v\lambda ij\exists n\exists s_0...s_n[i=s_0 \wedge j=s_n \wedge s_0[u := u\oplus 1]{*}s_n \wedge V(u)(s_n) = 0 \wedge \forall k{:}0\leq k<n\ V(u)(s_k) \neq 0 \wedge Agent(V(v)(s_0)) \wedge Patient(V(v')(s_0)) \wedge Sub(u)(V(v')(s_0))]$

Recall that the store u represents the mass of the object that is eaten. Yet so far this information is not present in (15). In order to make this information explicit, I will write Mass(u)(s) instead of V(u)(s). This yields (16).

(16) $\lambda v'\lambda v\lambda ij\exists n\exists s_0...s_n[i=s_0 \wedge j=s_n \wedge s_0[u := u\oplus 1]{*}s_n \wedge Mass(u)(s_n) = 0 \wedge \forall k{:}0\leq k<n\ Mass(u)(s_k) \neq 0 \wedge Agent(V(v)(s_0)) \wedge Patient(V(v')(s_0)) \wedge Sub(u)(V(v')(s_0))]$

The general interpretation of a transitive accomplishment-verb is given in (17).

(17) transitive accomplishment verb ⇝
$\lambda v'\lambda v\lambda ij\exists n\exists s_0...s_n[i=s_0 \wedge j=s_n \wedge s_0[u := u\oplus 1]{*}s_n \wedge V(u)(s_n) = \mathbf{b} \wedge \forall k{:}0\leq k<n\ V(u)(s_k) \neq \mathbf{b} \wedge \theta_1(V(v)(s_0)) \wedge \theta_2(V(v')(s_0)) \wedge Sub(u)(V(v')(s_0))]$

The formula in (17) consists of three parts the function of which can be described as follows.
(i) $[i=s_0 \wedge j=s_n \wedge s_0[u:=u\oplus 1]{*}s_n \wedge \forall k{:}0\leq k<n\ V(u)(s_k)\neq\mathbf{b} \wedge V(u)(s_n)=\mathbf{b}]$. This is the specification of the **while**-loop which has already been discussed in detail above. It constitutes the dynamic part of the representation. Notice that in this part the arguments $V(v)(s_0)$ and $V(v')(s_0)$ do not occur. The requirement that e.g. in the case of *eat* the value of $Mass(u)(s_0)$ must be the complete mass of the apple is handled by a presupposition.
(ii) *Agent(V(v)(s₀)) ∧ Patient(V(v')(s₀))*. In this part the roles of the arguments $V(v)(s_0)$ and $V(v')(s_0)$ are determined. It can be seen as the static part because there is no change with respect to the thematic relations. Note that by simply writing *Agent(V(v)(s₀)) ∧ Patient(V(v')(s₀))* we do not intend to claim that this is all what can or should be said about these thematic relations. It is certainly possible to analyse these relations further - to decompose them too, so to speak. Such a decompositional analysis can, for example, consist in relating the arguments $V(v)(s_0)$ and $V(v')(s_0)$ to each other or to specify properties of the thematic roles which can be used in an analysis of the (syntactic) argument structure of the verb. Furthermore it should be obvious that the representations given here for accomplishment-verbs do not capture all the meaning components of these verbs. Rather, what is represented is what is common to all accomplishment-verbs, at least from the perspective of aspectual phenomena. There are obviously differences in meaning which are not accounted for in the above representations. One possibility of at least indicating this lacuna would be to add a further predicate, say $\theta(V(v')(s_0))(V(v)(s_0))$, which stands as a kind of placeholder for the missing information. Yet, by incorporating the exact property which is changed by the denotation of the verb a first step to differentiate the various accomplishment-verbs has already been made.
(iii) *Sub(u)(V(v)(s₀))* or *Sub(u)(V(v')(s₀))*. This part relates the first two components to each other, that is, it establishes a relation between the dynamic and the static part. *Sub(u)(V(v)(s₀))* specifies one of the properties (or roles) of one of the arguments of the verb the value of which is changed by the assignment command occurring in the first component. Note that the change takes place with respect to this property and not with respect to the argument itself.

(3) The Interpretation of Activity-Verbs

By the initial assignment (environment or linguistic context) the domain S of the active variable u, i.e. the set of all values for which u is defined, is partitioned into two sets: $S = S_1 \cup S_2$ and $S_1 \cap S_2 = \emptyset$. S_1 is the set of all values that are excluded for u in the final state of a computation because if the value of u in that state would be any element from this set, no (possible) change would have occurred at all. For instance, in the case of eating an apple the value of u must not be identical to the total mass of the apple because this is the value that is initially assigned to u. Another example is given by a verb like *rise*, as in *The price of gold rose*. If the price of gold is initially, say, n dollar, then any value $n' \leq n$ is excluded as value in the final state for the active variable representing the price of gold. S_2, then, is the largest set of values relative to S_1 such that if the value of u in the final state is an element of this set, an execution of the program has brought about a (possible) change.

The question is what restrictions, if any, are imposed on S_2 with respect to the value of u in the final state of a computation. There are two principle cases to distinguish. Either no element of S_2 is excluded or some element(s) from this set is (are) excluded as possible value(s) (in the final state). In the first case the value of u can be said to be unrestricted with respect to S_2. In the second case, on the other hand, the value of u must be an element of a proper subset S_2' of S_2.

Recall that the basic element of the programs representing accomplishment-verbs is an assignment command with respect to the active variable, say u. Now suppose this assignment command is executed. In the first case described above the postcondition will already be satisfied after this execution because the value of u is an element of S_2. But this condition will also hold if the assignment command is executed more than once. More generally, it will be satisfied after an arbitrary number of executions of this program. In the second case this need not be true. As the value of u is required to be an element of a proper subset of S_2, it can happen that this condition holds only after the assignment command has been executed several times. It follows that there is the possibility that the program terminates in a state s without it being the case that the postcondition ϕ holds there, although the value of u will (necessarily) be an element of S_2. Take, for instance, the simulation of the while-loop in QDL that has been used to represent accomplishment-verbs: $(b?;\pi)^*$; $(\neg b)?$. The postcondition ϕ is that $\neg b$ holds. Now there are executions such that the program terminates in a state s but $\neg b$ does not hold in s. This always happens if the loop $(b?;\pi)^*$ is not executed 'long enough', so to speak.

This example shows that the interpretation of accomplishment-verbs is of the second kind. They impose stronger conditions on the value of u upon termination than simply being an element of the largest admissible set, i.e. S_2. To give an example, in the case of *eat* the condition is that the value of $\text{Mass}(u)(s_n)$ be 0 and not any value that is different from the one in the input state, i.e., from the total mass of the object that is eaten. Let us next analyse a verb like *push*.

(18) John pushed the cart.

In this case the relevant property concerns the object and its location (or position). The assignment command $u := u \oplus 1$ can therefore be interpreted as modelling a change of location according to a particular measure, for instance, one in yards or meters. The value of u in a given state then consists of the path that corresponds to the distance already traversed, and the value in the output state is the path that results if the path in the input is extended by

a single step. In this respect there is thus no difference between *push* and *cross*: both denote a change with respect to the location of an object denoted by an argument of the verb.

To which of the two cases does the restriction on u, the active variable representing the location of the object, belong? Obviously, *push* does not behave in the same way that, for example, *eat* does. Someone has pushed the cart if the cart has changed location, irrespective of the particular location of the cart. Thus, a verb like *push* belongs to the first kind. It does not put any further conditions on the value of the active variable than that it be an element of S_2. In terms of the constant V this amounts to the requirement $V(u)(i) \neq V(u)(j)$. It follows that this condition is already satisfied if the assignment command is executed only once. On the other hand, this condition will also be satisfied if the assignment command is executed further times. In effect, this condition will be satisfied after an arbitrary (non-empty) number of executions of this command.[5] For this reason it is not possible to interpret a verb like *push* in the same way as an accomplishment-verb, namely, as a **while**-loop.

(19) While the location has not yet changed continue pushing.

An interpretation based on (19) would yield only 'minimal' pushings in the sense that they are terminated as soon as the object, say, the cart, has moved at all. But this is not what one wants. According to what has been said above, one should rather have something like (20).

(20) $(u := u \oplus 1; \phi?)^*$

where ϕ is the condition that the value of u is different from the one in the initial state. According to (20), a verb like *push* denotes a kind of iteration. What gets iterated is the assignment command $u := u \oplus 1$ plus a test with respect to the relevant condition. Compare this program with a **while**-loop that is used to represent an accomplishment-verb.

(21) $((\neg\phi)? ; u := u \oplus 1)^* ; \phi?$

Note that both programs can generate extended execution sequences in the sense that the latter can consist of more than two states. This means the output-state need not be reached instantaneously but there can be intermediate states in between the input- and the output-state. In the case of a **while**-loop of the form (21) intermediate states are possible only if the condition ϕ is not guaranteed to hold after a single execution of the assigment command $u := u \oplus 1$. But this, in turn, can be the case only if the condition on the value of the active variable u in the output-state is stronger than to simply be an element of S_2, i.e., the maximal set of admissible values for a change to occur. Consequently, in order that intermediate states are possible, the set S_2 must be partitioned into two subsets S'_2 and S''_2 such that $[\![V(u)(s_n)]\!] \in S'_2$ must hold whereas $[\![V(u)(s_n)]\!] \in S''_2$ is excluded . The test $(\neg\phi)?$, then, corresponds to the case $[\![V(u)(s_i)]\!] \in S''_2$, $0<i<n$, whereas the final test $\phi?$ corresponds to $[\![V(u)(s_n)]\!] \in S'_2$. Thus, the two incompatible tests $(\neg\phi)?$ and $\phi?$ can both succeed only if there is a partition of S_2 into two sets, that is, if the condition on the value of u is not the weakest possible one. In the case of an activity-verb like *push*, on the other hand, there is no such partition of S_2 and, consequently, no strengthening of the condition on u takes place. It is just the weakest

[5] For an analysis of the problem of cyclic movements where one comes back to the original location see Naumann(95), chapter 6.

possible one. From this it follows that testing for $\neg\phi$, i.e., the negation of the requirement, will always fail. Taking the precondition into account, the translation of (20) into Type Logic yields (22).

(22) $\lambda v'\lambda v\lambda i\lambda j\exists n\exists s_0...s_n[i=s_0 \wedge j=s_n \wedge i[u := u \oplus 1]^*j \wedge Location(u)(s_0) = \mathbf{b} \wedge \forall k\ 0<k\leq n: \exists p[Location(u)(s_k) = p \wedge p \neq \mathbf{b}] \wedge \theta_1(V(v)(s_0)) \wedge \theta_2(V(v')(s_0)) \wedge Sub(u)(V(v')(s_0))]$

If there are zero executions of the iteration command, i.e., in case n = 0, the condition ϕ is not satisfied: in the output-state the value of u is not different from the one in the initial-state because the two states are identical. This is therefore not a possible execution of the program. Furthermore, note that the iteration-command that characterizes the interpretation of an activity-verb is also part of the representation of an accomplishment-verb. This will be important when one analyses directional PPs like *to the station* and modified VPs like *push the cart to the station* or *walk to the station* (see Naumann(95), chapter 7)

The result we have arrived at is that accomplishment-verbs like *eat* or *cross* and activity-verbs like *push* or *walk* are alike in the sense that both types of verbs express a change that can be temporally extended but need not be. At the formal level this similarity is reflected by the fact that in the representations of both kinds of verbs an assignment command occurs inside the iteration operator *. The difference between the two kinds shows up in the way in which the set of admissible values for u in the final state of a computation the obtaining of one of which guarantees that a change has occurred at all is constraint. Whereas accomplishment-verbs put further restrictions on S_2, this is not the case for activity-verbs. This difference is reflected in the kind of boolean condition that is used. Yet, more importantly, this difference has consequences for the control structure that is used. In the case of accomplishment-verbs the control-structure that is appropriate, i.e., which conforms to the intuitions, is that of a **while**-loop. Activity-verbs, on the other hand, are characterized by a (simple) iteration. The difference in control-structures can be said to be a consequence of the difference in the way the value of u in the final state is restricted in the following sense. Whenever the set S_2 is not further restricted, one is not interested in the *minimal* change for which this condition holds. Rather, what is important are arbitrary (finite) sequences for which the condition is satisfied. In case there are further conditions on the set S_2 the situation is different. Then it is possible to determine the *minimal* change that satisfies the requirement on u. This means the final state is the *first* state for which the condition on the value of u is satisfied.

So far the difference between activity-verbs and accomplishment-verbs has been characterized in terms of the programs that are used to interpret them. Another possibility of analysing the difference consists in comparing the properties these programs have instead of the programs themselves. The formalism in which these properties will be specified is Temporal Logic (TL). A property P of a program π is said to be specified by a TL-formula ϕ just in case ϕ holds for each computation sequence σ of π, where a sequence σ satisfies a formula ϕ if $s_0 \vDash \phi$ holds, i.e., if ϕ holds in the initial state of σ. There are different hierarchies of properties of programs that can be defined. A very basic one is based on the distinction between invariance (safety)-properties and non-invariant-properties, where the latter class consists of liveness- and precedence-properties. (for further details, see Naumann(96))

An invariance-property can be expressed by a formula of the form $\Box\phi$, where $\Box$ is the reflexive closure of the **G**-operator: $\Box\phi = \phi \wedge \mathbf{G}\phi$. The simplest way of specifying a liveness-property is by means of a formula of the form $\Diamond\phi$, where $\Diamond\phi = \neg\Box\neg\phi$. Precedence-

properties, finally, can be determined as combinations of safety- and liveness-formulas or, for instance, by means of an **Until**-formula **Until**(ϕ,ψ). The definition of this latter operator is given in (23).

(23) $s \vDash$ **Until**(ϕ,ψ) iff $\exists s'>s$: $s' \vDash \phi$ and for all s'' such that $s<s''<s'$: $s'' \vDash \psi$

All three types of properties can be generalized in different ways, as we will see below. The iteration-program interpreting an activity-verb can basically be characterized by a (generalized) invariance-property. To see this consider the two principle types of execution sequences generated by this type of program. (In what follows only the boolean condition ϕ is shown.)

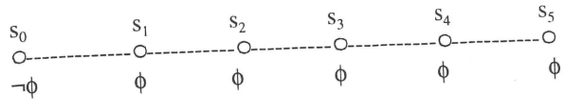


Figure 2

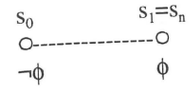


Figure 3

In both cases ϕ holds for all states except the initial one. This distribution of ϕ can be expressed by means of the generalized invariance-formula $\mathbf{G}\phi$:

(24) $s_0 \vDash \mathbf{G}\phi$

This formula expresses a generalized invariance-property because ϕ has to always hold in certain states of a computation. In the present case these states are all non-initial states. Using a particular predicate 'at s' to express that the control flow is at state s, (24) can also be expressed by the formula $\Box$[at $s_1 \vee ... \vee$ at s_n -> ϕ]. This generalized invariance-property reflects the fact that the postcondition ϕ holds after each execution of the underlying assignment command.

For a **while**-loop on the other hand the postcondition ϕ is satisfied only in the output-state s_n. The two principle types of execution sequence are given in Figures 4 and 5.

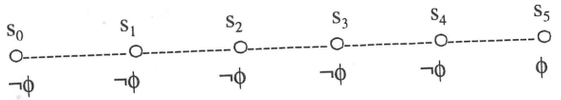


Figure 4

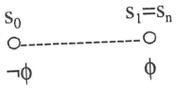


Figure 5

This uniqueness character can be expressed by means of the **Until**-operator.

(25) $s_0 \models \mathbf{Until}(\phi,\neg\phi)$[6]

As the underlying frames are linear, having both a first and a final state, the following formula can also be used.

(26) $s_0 \models \mathbf{F}(\mathbf{G}\bot \wedge \phi \wedge \mathbf{H}\neg\phi)$

Here $\mathbf{G}\bot$ is a formula that holds only at the final state of a computation. (26) then says that ϕ holds in the final state ($\mathbf{G}\bot \wedge \phi$) but that ϕ holds there for the first time ($\mathbf{H}\neg\phi$).

From the above characterization it follows that programs interpreting activity-verbs are prefix-closed in the following sense. Let $\sigma = s_0...s_n$ be a computation that belongs to the denotation of a program π. Then also each sequence $\sigma' = s_0...s_k$ for $0<k\leq n$ belongs to the denotation of π. (The minimal prefix $\sigma'=s_0$ must be excluded because in this case nothing has happened at all.)

(4) A Comparison to Event-based Approaches

Recall that in an event-based approach the difference between a verb like *eat* and verbs like *push* and *see* is analyzed in terms of the notion of graduality of one thematic role (incremental Theme vs. non-incremental Theme). Whereas the indirect object of *eat* is involved in an incremental way, the indirect object of *see* and *push* are involved in one go. For instance, the whole cart is pushed at once and a house is seen at once.[7] But it is possible to eat something in one go as well. Take elephants as an example. They can eat a complete apple in one go and human beings can do the same with peanuts. In cases like these the event is not gradual in the sense that proper parts of the apple (peanut) are mapped to proper parts of the event but rather all parts of the apple (peanut) are mapped to the whole event which doesn't have any proper temporal parts at all. Thus the situation is similar to that of seeing a house or pushing a cart (at least if one assumes that the latter two are only minimal in the sense that the event has no proper temporal parts as well). The problem at hand concerns execution sequence of the type depicted in Figure 5 above. Examples like these show that graduality is not satisfied for all eating-events, or, more generally, for all events that have an Incremental Theme. It is at least in principle always possible to subject this Theme to the event in one go. It follows that if all eating events were of the 'in one go'-variety described above, they should be classified as durative and not as terminative. Therefore graduality should not be taken to be a necessary condition for terminative readings (modulo the referential properties of the argument NPs).

It is nevertheless possible to formulate the difference between *eat* and *push* in terms of properties of the thematic relation assigned to the internal argument. The thematic role θ assigned to the internal argument of verbs like *eat* can be characterized by the following property. (compare Eberle(95) who argues along similar lines.)

[6] Strictly speaking, the condition that ϕ does not hold at s_0 has to be included. Yet this condition follows from the presupposition (precondition) according to which ϕ never holds at s_0.

[7] At least under one reading of *see*. There is another reading according to which a house can be seen gradually.

(27) $\forall e,e',x[P(e) \wedge \theta(e,x) \wedge e' \sqsubset e \wedge P(e') \Rightarrow \exists x'[x' \sqsubset x \wedge \theta(e',x')]]$

According to (27), graduality is not necessarily involved in events denoted by a verb like *eat* or *cross*. The antecedent can be false because there is no proper subevent e' of e such that $e' \sqsubset e$ does not hold. What is characteristic of this role rather is the fact that if there are proper parts e' of an event then θ will relate e' not to x, i.e., the object to which e is related, but to a proper (material) part x' of it. The situation is different for thematic relations like Agent or the Patient-role assigned to the internal argument of *push*. Here to each proper subevent e' of e the thematic role assigns the same object. The role is therefore constant.

(28) $\forall e,e',x[P(e) \wedge \theta(e,x) \wedge e' \sqsubset e \wedge P(e') \Rightarrow \theta(e',x)]$

A consequence of defining the characteristic property of the relevant thematic relations in this way is that the influence of plural argument NPs must be analyzed in a way different from what is usually assumed in event-based approaches. Note that the definitions in (27) and (28) refer to the objects denoted by the argument-NP and therefore to the ontological level and that no mention is made of any semantic properties the argument NPs may have. This leads to a problem in the case of constant roles. Consider the following two sentences.

(29) a. John pushed three carts.
b. John pushed carts.

In both cases it is possible that John pushed the carts involved successively, that is, one after the other. This means that for a proper subevent e' of the (complete) pushing the object assigned to e' can be a proper part (or a proper subset) of the object assigned to e by θ, contrary to what is required by (28). Therefore the constancy that is required by (28) only holds for an atomic object but not (necessarily) for plural ones. For that reason the definitions in (27) and (28) should include the condition that x is an atom (or a singleton).

The relationship between graduality and terminativity/durativity in event-based approaches makes it very difficult, if not impossible, to account for the aspectual properties of verbs like *hit* or *knock*. Such verbs are terminative. There are two principle strategies one can pursue in order to account for this behaviour. Either one analyses them as a separate class, besides accomplishments and activities, or one just describes them as accomplishments. In the former case one has to explain their aspectual properties in a way different from that used for the other two classes. This seems to be very difficult. In the latter case the thematic role assigned to an object denoted by an argument-NP can be gradual only in a trivial sense. Consider for example *hit*. If there are any proper subevents e' of a hitting-event e at all, i.e., if hit(e') holds, they too involve the object as a whole and not a proper part of it. But this means that the internal argument-NP should get a constant role and not a gradual one. A similar argument holds for verbs like *reach* or *win* that are often classified as achievements. In the case of *reach* e.g. there will normally be a movement of the subject before the reaching takes place. But this preparatory phase, as it may be called, cannot itself be described as an event of reaching. It follows that at the level of verbs terminativity does not presuppose the existence of a gradual role.

In order to be able to accomodate the possibility of configurations as that shown in Figure 5 one has to distinguish between the possibility of generating execution sequences with intermediate states during which a change occurs and the question of whether the interpretation of a verb determines a terminating condition, where this condition corresponds

to a non-invariant property of the program. The former possibility is implemented by means of the Kleene-star *. Its occurrence is independent of the distinction between terminativity and durativity. As will be shown in section (5), the difference between programs of which the Kleene-star is part and those for which this does not hold can be used to distinguish accomplishment-verbs from point-verbs. The distinction between terminativity and durativity has to do rather with the distinction between invariance-properties and non-invariant ones, a distinction that is formally implemented by means of the boolean condition ϕ.

The constancy that is claimed for the thematic roles of activity-verbs like *push* in event-based approaches corresponds to the fact that the boolean condition ϕ is a generalized invariant of the program interpreting this class. As was shown above, from this it does not follow that these verbs are not gradual. Graduality, on the other hand, is implemented by means of the Kleene-star and an assignment-command. When taken together these two programs allow for the possibility of (temporally) extended changes. Programs in which these two programs are combined differ both from programs that lack any change component in form of the assignment-command (i.e., programs that are basically tests) and programs in which the Kleene-star does not occur. The difference between programs interpreting accomplishment-verbs and those interpreting activity-verbs comes in with the boolean condition. It is this component that underlies the distinction between terminativity and durativity.

The above considerations show one advantage of interpreting verbs as programs over an interpretation that is based on the notion of an event. By using program-concepts it is possible to interpret verbs as kinds of complex programs. The various component programs can be used to represent different aspects of the dynamic properties of the verbs that are aspectually relevant. Thus an assignment command expresses the notion of change, whereas the Kleene-star models the possibility of extended changes. The boolean condition, finally, serves two different purposes. First it determines the condition on the active variable in the output-state (the postcondition) and second it is specified whether this condition only holds in this state or also in other states of a computation.

Whereas the first two types of programs, i.e., the assignment command and the Kleene-star, can be interpreted straightforwardly in a realistic way, this does not hold for the boolean condition. It rather reflects more purely linguistic (semantic) information that does not necessarily correspond to any features of the situation described by the expression. The various programs are independent of each other in the sense that it is possible to use one without any of the remaining ones. It is this independence that makes it possible to distinguish semantic properties of verbs according to the different types of programs. There can be verbs that are alike in the use that is made of some types of programs, yet which differ with respect to other types.

(5) The Interpretation of other Classes of Verbs

(5.1) The Interpretation of Point-Verbs

(30) a. John knocked at the door in five minutes.
b. John knocked at the door for five minutes => iteration of knockings
c. John knocked at the door in five minutes ≠> John was knocking at the door during this period
d. John was hitting Mary ≠> John has hit Mary
e. John knocked at the door at noon.

As the examples in (30a) and (30b) show, point-verbs are terminative: they can be modified by *in*-adverbials, whereas the modification by *for*-adverbials leads to an iterative reading. But there are clear differences to the behaviour of accomplishment-verbs. The example in (30c) shows that from 'John VP-ed in n time' it does not follow that John was VP-ing during the period denoted by the time adverbial. The interpretation of (30c) is that the knocking took place at the end of the five minutes period and that there are no 'knocking-activities' during this period. The case of an accomplishment-verb is different. Here the inference from, say, *John ate an apple in two minutes* to *John was eating an apple during this period* is valid (if possible pauses count as eating). The example in (30d) shows that the imperfective paradox holds for point-verbs as well. But here too there is a difference with respect to the interpretation of the progressive when compared to accomplishment-verbs. *John was hitting Mary* can easily be interpreted as referring to a single hitting by John which is part of a sequence of several hittings. Such an interpretation is difficult to get in the case of an accomplishment-verb: without a special context *John was crossing the street* can only mean that there was exactly one crossing intended. The example in (30e) finally shows that point-verbs allow modification with point-adverbials in contrast to accomplishment-verbs for which such modifications are at least problematic. The difference to activity-verbs consists in the fact that the latter can only be interpreted inchoatively with such a modification. For instance, *John ran at noon* can only mean that John started to run at noon but not that a whole running period took place at that time.

How shall point-verbs be interpreted? Above in section (4) it was shown that execution sequences like that shown in Figure 5 pose difficulties for approaches which essentially link terminativity to graduality. The above data can be interpreted as showing that the programs denoted by point-verbs are generally of the type of programs corresponding to the execution sequence of Figure 5. (for more details see Naumann(95), chapters 6,7) The difference to the interpretation of accomplishment-verbs then has to do with the cardinality of the domain of the active variable. In the case of accomplishment-verbs this domain can in principle be infinite. For point-verbs, on the other hand, it will be assumed that this domain consists of exactly two elements: $S = \{v_1, v_2\}$. It follows that S can only be partitioned in one way into two non-empty sets S_1 and S_2: $\{\{v_1\},\{v_2\}\}$, if v_1 and v_2 are the two elements. Recall that the minimal requirement for the occurrence of a change is that the value of the active variable u in the final state is different from the one in the initial state: $[\![V(u)(i)]\!] \neq [\![V(u)(j)]\!]$. In the case of a two-element domain there is exactly one way of satisfying this condition, namely, by assigning the other value to u, i.e., the value that does not hold initially. It follows that no non-relative specification is necessary (nor is it possible). The same holds for the iteration command because the only possible change already occurs by executing the assignment command once. For that reason the two values can be identified with the truth values $\top$ and $\bot$: $S = \{\top,\bot\}$. The assignment command underlying the interpretation of point-verbs can then be defined as $[u := \neg u]$.

(31) a. $[u := \neg u] = \lambda ij[V(u)(j) = \neg V(u)(i) \wedge \forall v[v \neq u \rightarrow V(v)(i) = V(v)(j)]]$
b. $\lambda v'\lambda v\lambda ij[i[u := \neg u]j \wedge V(u)(j) = \neg V(u)(i) \wedge \theta_1(V(v)(s_0)) \wedge \theta_2(V(v')(s_0)) \wedge Sub(u)(V(v')(s_0))]$

In (31a) the translation of the assignment command is given (where it is assumed that $\top = \neg\bot$), whereas (31b) contains the interpretation of a transitive point-verb like *hit* or *beat*.

On the other hand, it is possible to represent a point-verb by means of a **while**-loop $((\neg\phi)? ; \pi)^* ; \phi?$, where ϕ is the minimal condition for a change.

(32) $\lambda v'\lambda v\lambda ij\exists n\exists s_0...s_n[i=s_0 \wedge j=s_n \wedge s_0[u := \neg u]^*s_n \wedge V(u)(s_n) = \neg V(u)(s_0) \wedge \forall k:0\leq k<n$
$V(u)(s_k) = V(u)(s_0) \wedge \theta_1(V(v)(s_0)) \wedge \theta_2(V(v')(s_0)) \wedge Sub(u)(V(v')(s_0))]$

As the test $(\neg\phi)$?, i.e., $V(u)(s_k) = V(u)(s_0)$, can succeed exactly once (in s_0) before the condition ϕ holds for the first time (in s_1), the computations σ consist of only two states: $\sigma = s_0s_1$ with $s_1 = s_n$. The iteration command is, so to speak, superfluous because there can be no non-trivial iteration of its argument due to the fact that it can only successfully be executed once. The execution sequences that are generated correspond to the limiting execution sequences generated by programs interpreting accomplishment-verbs, that is, to sequences of the type depicted in Figure 5. (Sequences of more than one beating or hitting are treated as a form of iteration; for details see Naumann(95), chapter 7)

(5.2) The Interpretation of stative Verbs

Sentences with stative verbs are durative as the following examples show.

(33) a. John was ill for two weeks/*in two weeks.
b. Children were ill for two weeks /*in two weeks.
c. John hated three yuppies for three months/*in three months.
d. John hated yuppies for three months/*in three months.

The referential properties of the argument NPs have no influence on the aspectual properties of the sentence and it is equally not possible to change the durative interpretation to a terminative one by means of a modification as in the case of verbs like *push* or *walk*. These facts can be accounted for by interpreting stative verbs as involving no change at all. This interpretation also accords with the intuition speakers have with respect to these verbs.

Formally this means that their interpretation does not involve an assignment command. From this it does not follow that they denote no programs at all but only that they must denote programs that do not change the input-state. They can therefore be interpreted as tests. The interpretation of the verb *hate* is given in (34).

(34) hate $\rightarrow$
$\lambda v'\lambda v\lambda ij[i=j \wedge hate(V(v')(j))(V(v)(j))]$

This type of interpretation is exactly the interpretation that is given to all transitive verbs by Muskens in his 'Logic of Change'. Note that no decomposition is needed, but it is of course possible to use such a decompositional analysis too, if one wishes. A possible example is given in (35).

(35) transitive stative verb
$\lambda v'\lambda v\lambda ij[i=j \wedge V(u)(j) = \mathbf{b} \wedge \theta_1(V(v')(j)) \wedge \theta_2(V(v)(j))]$

In (35) u denotes some property which must have a particular value **b** in the output-state j, which is identical to the input-state such that in effect it is tested whether store u has value **b** in the output-state. It may seem that the above definition is inadequate because it allows a state to hold only minimally such that sentences like *John was ill for two weeks* cannot be accounted for. This objection is not correct. The interpretation that has been given so far

accounts only for what may be called the dynamic structure. This structure must be distinguished from the temporal structure which is treated separately. There will of course be constraints on the temporal structure that are induced by the dynamic component. For instance, one will require an axiom like $<\pi^+>\phi$ -> $<\mathbf{F}>\phi$ (where $\pi+ = \pi;\pi^*$) to hold which says that if ϕ holds after a program π is executed an arbitrary, non-zero number of times, then ϕ holds in the future, that is, if a state at which ϕ holds is reachable by executing a program π a finite number of times, then this state will be later than the initial state.

(6) The Classification of Verbs

Let me finally discuss what classification of verbs one arrives at according to the present approach. The important point to note is that there is no unique classification but at least two. The first classification is mainly based on two oppositions: test command/assignment command and invariance-properties/non-invariance-properties. The first opposition underlies the distinction between stative-verbs and non-stative ones. This distinction is based on the intuition that the former class does not involve the notion of change whereas the latter class essentially involves this notion in its interpretation. The second opposition characterizes the distinction between what has traditionally been called processes and events. Intuitively, processes (activities) denote a kind of indefinite change whereas events denote a kind of definite change. Formally, this distinction is explained as a difference with respect to the boolean conditions that are determined and that specify different properties of programs: invariance-properties versus liveness- or precedence-properties. The difference between point-verbs and the other two classes that are classified as events is then based on the fact whether the iteration command is part of the representation or not. (For the interpretation of achievement-verbs see Naumann(95), chapter 7)

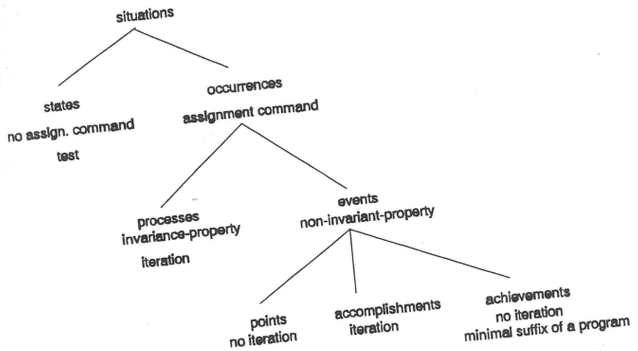


Figure 6

The second way of classifying the different classes of verbs is based on the notion of prefix-closedness, or, to be precise, a generalization of this notion as defined above. A program π is generalized prefix-closed iff $s_0 \models \Box\phi \vee \mathbf{G}\phi$. As was shown above, programs interpreting process-verbs satisfy the second disjunct, $\mathbf{G}\phi$. The first disjunct is satisfied by tests because the input-state is identical to the output-state such that ϕ holds at the only state s_0 (=i=j) which means that $\Box\phi$ holds in this state.

(32) $\lambda v'\lambda v\lambda ij\exists n\exists s_0...s_n[i=s_0 \wedge j=s_n \wedge s_0[u := \neg u]^*s_n \wedge V(u)(s_n) = \neg V(u)(s_0) \wedge \forall k:0\leq k<n$
$V(u)(s_k) = V(u)(s_0) \wedge \theta_1(V(v)(s_0)) \wedge \theta_2(V(v')(s_0)) \wedge Sub(u)(V(v')(s_0))]$

As the test $(\neg\phi)?$, i.e., $V(u)(s_k) = V(u)(s_0)$, can succeed exactly once (in s_0) before the condition ϕ holds for the first time (in s_1), the computations σ consist of only two states: $\sigma = s_0s_1$ with $s_1 = s_n$. The iteration command is, so to speak, superfluous because there can be no non-trivial iteration of its argument due to the fact that it can only successfully be executed once. The execution sequences that are generated correspond to the limiting execution sequences generated by programs interpreting accomplishment-verbs, that is, to sequences of the type depicted in Figure 5. (Sequences of more than one beating or hitting are treated as a form of iteration; for details see Naumann(95), chapter 7)

(5.2) The Interpretation of stative Verbs

Sentences with stative verbs are durative as the following examples show.

(33) a. John was ill for two weeks/*in two weeks.
b. Children were ill for two weeks /*in two weeks.
c. John hated three yuppies for three months/*in three months.
d. John hated yuppies for three months/*in three months.

The referential properties of the argument NPs have no influence on the aspectual properties of the sentence and it is equally not possible to change the durative interpretation to a terminative one by means of a modification as in the case of verbs like *push* or *walk*. These facts can be accounted for by interpreting stative verbs as involving no change at all. This interpretation also accords with the intuition speakers have with respect to these verbs.

Formally this means that their interpretation does not involve an assignment command. From this it does not follow that they denote no programs at all but only that they must denote programs that do not change the input-state. They can therefore be interpreted as tests. The interpretation of the verb *hate* is given in (34).

(34) hate ⇀
$\lambda v'\lambda v\lambda ij[i=j \wedge hate(V(v')(j))(V(v)(j))]$

This type of interpretation is exactly the interpretation that is given to all transitive verbs by Muskens in his 'Logic of Change'. Note that no decomposition is needed, but it is of course possible to use such a decompositional analysis too, if one wishes. A possible example is given in (35).

(35) transitive stative verb
$\lambda v'\lambda v\lambda ij[i=j \wedge V(u)(j) = \mathbf{b} \wedge \theta_1(V(v')(j)) \wedge \theta_2(V(v)(j))]$

In (35) u denotes some property which must have a particular value **b** in the output-state j, which is identical to the input-state such that in effect it is tested whether store u has value **b** in the output-state. It may seem that the above definition is inadequate because it allows a state to hold only minimally such that sentences like *John was ill for two weeks* cannot be accounted for. This objection is not correct. The interpretation that has been given so far

accounts only for what may be called the dynamic structure. This structure must be distinguished from the temporal structure which is treated separately. There will of course be constraints on the temporal structure that are induced by the dynamic component. For instance, one will require an axiom like $<\pi^+>\phi$ -> $<\mathbf{F}>\phi$ (where $\pi+ = \pi;\pi^*$) to hold which says that if ϕ holds after a program π is executed an arbitrary, non-zero number of times, then ϕ holds in the future, that is, if a state at which ϕ holds is reachable by executing a program π a finite number of times, then this state will be later than the initial state.

(6) The Classification of Verbs

Let me finally discuss what classification of verbs one arrives at according to the present approach. The important point to note is that there is no unique classification but at least two. The first classification is mainly based on two oppositions: test command/assignment command and invariance-properties/non-invariance-properties. The first opposition underlies the distinction between stative-verbs and non-stative ones. This distinction is based on the intuition that the former class does not involve the notion of change whereas the latter class essentially involves this notion in its interpretation. The second opposition characterizes the distinction between what has traditionally been called processes and events. Intuitively, processes (activities) denote a kind of indefinite change whereas events denote a kind of definite change. Formally, this distinction is explained as a difference with respect to the boolean conditions that are determined and that specify different properties of programs: invariance-properties versus liveness- or precedence-properties. The difference between point-verbs and the other two classes that are classified as events is then based on the fact whether the iteration command is part of the representation or not. (For the interpretation of achievement-verbs see Naumann(95), chapter 7)

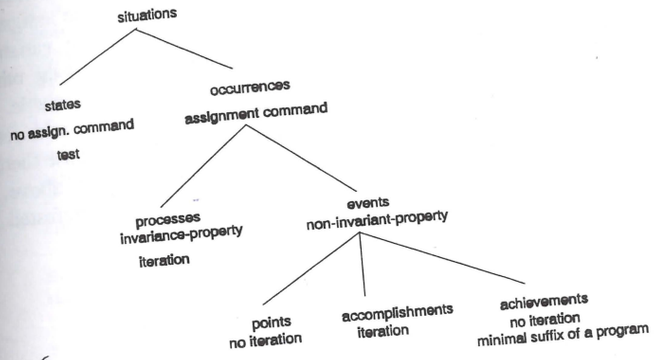


Figure 6

The second way of classifying the different classes of verbs is based on the notion of prefix-closedness, or, to be precise, a generalization of this notion as defined above. A program π is generalized prefix-closed iff $s_0 \models \Box\phi \vee \mathbf{G}\phi$. As was shown above, programs interpreting process-verbs satisfy the second disjunct, $\mathbf{G}\phi$. The first disjunct is satisfied by tests because the input-state is identical to the output-state such that ϕ holds at the only state s_0 (=i=j) which means that $\Box\phi$ holds in this state.

Obviously this notion of generalized prefix-closedness is closely related to the notion of divisivity used in event-based approaches to characterize (at least partly) stative verbs and process-verbs. One problem with using this notion for process-verbs is that they satisfy it only down to a certain degree (or down to a certain threshold). For instance, to waltz takes at least three steps and to walk takes at least one step. On the present account this difference is easily explained. Process-verbs are interpreted by non-test-like programs the execution of which involves at least one transition to an output-state that is different from the input-state. In the case of waltzing, e.g., a single transition must correspond to at least three steps. This just reflects the fact that such programs are modelling a change. But a change can only have occurred if the underlying assignment command has been executed at least once.

What is common to both kinds of programs is the property that if the underlying basic program, i.e., the test or the assignment command, is (successfully) executed once the postcondition ϕ is already satisfied in the output-state. As was shown above, this corresponds to the fact that no progress can be made towards the obtaining of ϕ in the sense that there can be intermediate states at which ϕ does not (yet) hold. Viewed from this perspective, point-verbs also belong to this class, as it should already be obvious from the fact that $s_0 \models \mathbf{G}\phi$ holds for this class. It is just this possibility of progress that characterizes programs interpreting accomplishments-verbs and achievement-verbs, if in the case of the latter the whole program of which the single transition is a part is considered. Thus the second classification can be said to be based on the notion of progress which can formally be defined as the negation of prefix-closedness.

The reason why point-verbs can be classified together with stative-verbs and activity-verbs has to do with the way in which the boolean condition functions. As was mentioned in Section (5.1), in the case of point-verbs the boolean condition determines the postcondition only in a relative way, i.e., only in relation to the value of the active variable in the initial state. In this respect point-verbs are similar to activity-verbs and opposed to accomplishment-verbs. This similarity follows from the fact that in case there is no occurrence of the Kleene-star the distinction between (generalized) invariance-properties and non-invariant ones is in a certain sense neutralized. The program simply consists in the execution of the assignment-command. There is no longer any control-structure. Assigning to the active variable an arbitrary value that is different from the one in the initial state and assigning it the other of two possible values amounts in both cases to changing the value of the active variable. Thus the difference between point-verbs and activity-verbs reduces to the difference with respect to the cardinality of the domain of the active variable. Whereas in the former case there are only two values, the domain can be infinite in the latter one. As was shown above, this possibility results in a particular control-structure because one is no longer interested in a minimal change.

Bibliography

Eberle, K.: 1995, *The Influence of Plural NPs on Aktionsart in DRT*, manuscript, University of Stuttgart

Krifka, M.: 1992, Nominal reference, temporal constitution and thematic relations, in I. Sag et al. (eds.), *Lexical Matters*, CSLI Publications, Stanford, 29-53

Muskens, R.: 1994, A compositional discourse representation theory, in P. Dekker/M. Stokhof (eds.), *Proceedings 9th Amsterdam Colloquium*, ILLC/Department of Philosophy, University of Amsterdam, 467-486

Naumann, R.: 1995, *Aspectual Composition in Dynamic Logic*, Habilitationsschrift, University of Düsseldorf

Naumann, R.: 1996, *Temporal Logic as a Specification Logic for Aspect*, Unpublished manuscript, University of Düsseldorf

Verkuyl, H.J.: 1993, *A Theory of Aspectuality*, Cambridge University Press, Cambridge

Radical Underspecification[1]

Manfred Pinkal
Universität des Saarlandes

1 Three Levels of Semantic Underspecification

Natural language can be conceived of as a system of rules or principles that serve the purpose to assign semantic representations to utterances on the basis of a given amount of linguistic information, which can be read off the utterance, plus situational or contextual information in the widest sense. An utterance is vague if the semantic representation assigned to it does not completely determine its denotation. It is semantically underspecified, if the given linguistic and contextual input information is not sufficient to uniquely determine a semantic representation. Formalisms for underspecified semantic representation offer techniques for the compact description of the possible range of representations, as an alternative to the mere disjunction of readings. Their usefulness is evident, if there are methods available for direct further processing of underspecified representations. A typical case may be semantic transfer rules in Machine Translation, which in certain cases bring about appropriate translations by associating the underspecified source language representation with an equally underspecified target language representation. Another possible use for underspecified representation is in modelling the process of utterance interpretation: an utterance may, e.g., be processed by first interpreting it in complete absence of contextual information, which usually leads to an enormous degree of underspecification, and then reducing the range of possible readings by adding non-linguistic information of different kinds.

Different kinds or levels of underspecification can be distinguished according to their structural properties, which pose different problems for the design of an underspecification formalism. On a first level, we have cases of local underspecification in an otherwise uniquely given global semantic structure. Typical examples are lexical ambiguities (like *pen* and *light*), referential ambiguities, as they occur with anaphoric or deictic uses of pronouns and with definite descriptions, and cases of relational vagueness, for which the genitive attribute (like *John's book*) is a typical case. These cases, which we will refer to as *Level-1 Underspecification* in this paper, have been the subject of extensive discussion, particularly with respect to their compositionality properties on the level of denotations (cf. Pinkal 1985, Pinkal 1995).

Scope ambiguities, on the other hand, as the one given in Example (1), are the classical representatives of a second level of semantic underspecification: In these cases, the global semantic structure of the representation is underspecified. Other phenomena falling into this category are focus-ground distinctions, and the distinction between collective and distributive readings. Examples are (2) and (3), respectively.

(1) *Every student did not pay attention*
(2) *Only intelligent students are successful*
(3) *Three boys lifted four pianos*

These *Level-2 phenomena* have been intensively discussed in the last years. No full and adequate denotational semantics has been developed for them. The lack of a

1. The work reported here arose partly from experience with dialogue translation in the Verbmobil project (funded by BMBF, German Ministery for Education and Research) and partly from many discussions in the FraCaS project (funded by the EC in the LRE program). It was also inspired by work of Reinhard Muskens and Wojciech Skut. I am especially indepted to Michael Kohlhase for hints and helpful comments concerning technical aspects.

unique underlying structure of representations is a hard problem for a compositional denotational interpretation. However, there is a variety of similar attempts available, which treat underspecification on the representational level (cf. QLF, Alshawi and Crouch 1992, UDRT, Reyle 1993, LUD, Bos 1996, MRS Egg and Lebeth 1995, Poesio 1994, Muskens 1995, van Eijck 1996). The common idea in most of these approaches is to provide a partial description of semantic representations in terms of constraints using meta-variables. As an example, we give a Reyle-style scope-underspecified representation of sentence (1) in (4), which covers the two specific scope readings of (5). The representation uses so-called labels l_i as meta-variables, and an additional relation of semantic subordination as a relation between labels.

(4) $\{l_1 : \forall x(student(x) \rightarrow l_2), l_3 : \neg l_4, l_0 : pay\_attention(x),\ l_1 \leq l_T,\ l_2 \leq l_T,\ l_0 \leq l_2,\ l_0 \leq l_4\}$

(5) $\{\forall x(student(x) \rightarrow \neg pay\_attention(x)),$
$\neg\forall x(student(x) \rightarrow pay\_attention(x))\}$

On the level of representation, Level-1 cases can be treated as trivial subcases of Level 2. The pair of readings in (8) for Sentence (6) is described by (7), where the first constraint gives a complete structural description containing a meta-variable to represent the lexical ambiguity.

(6) *Bill owns a pen*

(7) $\exists x(X(x) \wedge own(b, x)), X \in \{pen_1, pen_2\}$

(8) $\exists x(pen_1(x) \wedge own(b, x)), \exists x(pen_2(x) \wedge own(b, x))$

There are additional underspecification phenomena, however, which cannot be subsumed under one of the two classes described above: Semantic underspecification induced by ambiguous, incoherent, or incomplete grammatical (i.e., extra-semantic) linguistic information. We give examples for the different subcases of these *Level-3 phenomena*, in the following, and will discuss their specific properties and problems afterwards.

Ambiguous syntactic information:

(9) *John saw the man with the telescope*

(10) *Die Linguistin zitiert das Buch*
$[_{NP\,nom\vee acc}$*Die Linguistin*$]$*zitiert*$[_{NP\,nom\vee acc}$*das Buch*$]$
The linguist cites the book / The book cites the linguist

(9) exemplifies the well-known attachment ambiguity. In (10), the grammatical function of the NP complements are underspecified, since subject and object can take the sentence-initial position in German, and the morphological information attached to the NPs *die Linguistin* and *das Buch* is not sufficient to identify case. This grammatical underspecification induces semantic underspecification with respect to the association of terms with argument positions.

Incoherent linguistic information:

(11) * *Den Logiker zitiert den Aufsatz*

(11) is a typical example in spoken-language processing: Speech recognition has mistakenly identified both articles as accusative forms, which leaves us in an information state that precisely corresponds to the information in example (10): We know that it is a transitive construction with a subject and a direct object, but we do not know which is which.

Incomplete linguistic information:

(12) _ *Logiker zitiert* _ *Aufsatz*

(13) ... *Logiker ... zitiert ... Aufsatz ...*

(12) indicates the case where the article positions could not be identified (with the same effect, again, as in (10) and (11)). (13) is an example for the much more

radical case where some single words have been recognized, but there is no explicit information at all available about the structure of the utterance, and about the role which the words play in that structure, which typically happens if an utterance is made in a noisy environment, or if keyword spotting techniques are used in an artificial speech recognition system.

It is evident that these *radical underspecification* phenomena are important. In particular, a treatment of cases like (11)-(13) is a necessary prerequisite for semantic processing of spoken input in any "real-world" application. Unfortunately, in sharp contrast to the variety of frameworks available for Level-1 and Level-2, Level-3 phenomena rarely have been taken notice of[2]. Also, there is no straightforward way of extending formalisms for scope underspecification to Level-3 cases. Unlike scope phenomena, these cases immediately require expressing semantic information of arbitrary degrees of underspecification: in the extreme case (13) we just know that the semantics of an identified subexpression contributes in some way or the other to the semantics of the full utterance, but we do not know for sure in which way it contributes. Also, Level-3 phenomena raise the question what the notion of compositional analysis might mean at all—even at the level of representation—given that there is no uniquely identifiable underlying syntactic structure for the utterance to be analyzed.

Our proposal is an attempt to solve the task of the representation of arbitrary underspecified semantic information. We will employ meta-variables that range over semantic operations, in addition to the meta-variables for object-language expressions used in the scope underspecification formalisms. The use of operator variables will also provide a solution to the task of semantic composition. Scope underspecification (as Level-2 phenomena in general) will turn out to be a special case of a much more general class of ambiguity cases.

In Section 2, we give a first illustration of how our higher-order constraint account works, looking at a simple example of syntactic ambiguity. In Section 3, we discuss which kinds of restrictions on our description language may be necessary or useful. Section 4 specifies an operational semantics for the constraint language in terms of Linear Unification, a modification of Huet's Higher-Order Unification Algorithm. In Section 5, we demonstrate the application of the formalism to syntactic and scope ambiguity examples. Section 6 concludes the paper with a discussion of further possible applications and required extensions of the formalism.

2 Underspecification with Higher-Order Constraints

Let us start with the simplest case. Take $X = walk'@j^*$ to be the fully specified semantic description of the object language representation $walk'(j^*)$ for the sentence *John walks*. The operation of functional application is explicitly represented by the (left-associative) operator @, taking object language functor and argument as its left and right argument, respectively. By means of meta-variables X and Y, we can form the underspecified descriptions $X = walk'@Y$ and $X = Z@j^*$, which are intended to specify the sets of all expressions of the form $walk'(\alpha)$, and $\beta(j^*)$, respectively, where α and β are some object-language expressions of appropriate type. Note that Y and Z are supposed to take expressions of different OL types as values, but both are first-order variables of the description language. Adding second-order variables C, C' to the description language also allows to underspecify the semantic operation, as in $X = C(walk', j^*)$ and $X = C'(walk')$. By substituting the operator variables C and C' with $\lambda XY.X@Y$ and $\lambda X.X@j^*$, respectively, we obtain $X = walk'@j^*$. However, $X = C'(walk')$ is supposed to cover not only the

2. In fact it has been an observation by Wojciech Skut which led my attention to the higher order character of Level-3 phenomena the first time. Cf. Skut 1995

$walk'(\alpha)$ cases, but in addition those cases where $walk'$ takes the argument position of a functional application, as in $every'(man')(walk')$, or appears arbitrarily deeply embedded either in the functor part of a function application, or in its argument part, or in an argument of any other semantic operation.

As object language, we will use the language of extensional type theory throughout this paper, for the sake of simplicity. Note, however, that our approach is not bound to any specific semantic representation language. We could have used any compositional semantic formalism instead, e.g., CDRT (Muskens 1993), λ-DRT (Kuschert 1996), or a compositional version of Situation Semantics (Cooper 1993).

We will call our description language "USDL" (Underspecified Semantic Description Language). Its inventory is given in (14).

(14) Symbols of USDL are:

(i) Logical constants: λ, $=$, and generalized conjunction, which is used to form finite conjunctive constraint sets (the latter being written in the usual set notation, without an explicit operator symbol).

(ii) A set of non-logical individual constants CON_0 denoting constants and variables of the object language, and sets of n-place function constants CON_n, denoting object-language operations (e.g., $@ \in CON_2$ for function application, $L_v \in CON_1$ for every OL variable v, denoting abstraction over v).

(iii) A set of first-order variables VAR_0 (written as X, Y, Z,...) ranging over object language expressions, as well as sets of n-place second-order variables VAR_n for every $n \in N$ (written as C_1, C_2,...).

If we take USDL to be a reduced version of the language of typed λ-calculus, the notion of a well-formed expression follows immediately from (14). We designed the language as lean as possible - no Boolean operations except conjunction, no quantifiers - to keep the formalism simple at first. Although it is rather certain that we will need extensions of USDL for more complex examples (see Section 6 below), we will stick to the inventory given in (14) in this paper.

We illustrate the way we employ the description language with (15), a simple example of syntactically induced underspecification. (15) has the two possible syntactic analyses in (16), which lead to the type-theoretic representations given in (17).

(15) *former professor in Amsterdam*

(16) (a) $[_{N'}\ [_{A}\ former]\ [_{N'}\ [_{N}\ professor\ ]\ [_{PP}\ in\ Amsterdam]]]$

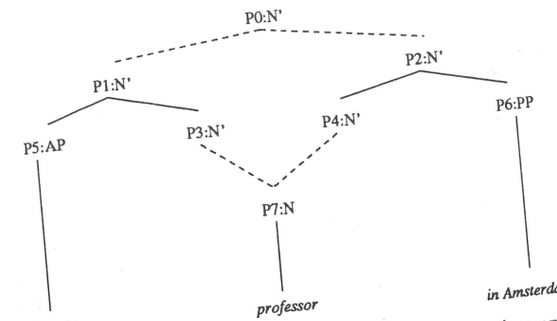
(b) $[_{N'}\ [_{N'}\ [_{A}\ former]\ [_{N}\ professor\ ]]\ [_{PP}\ in\ Amsterdam]]$

(17) (a) $former'@(in\_Amsterdam'@professor')$

(b) $in\_Amsterdam'@(former'@professor')$

In the following, we will demonstrate how underspecified semantic descriptions are assigned to underspecified syntactic representations. As a prerequisite, we need a method for compact representation of syntactic ambiguity. We use the general dominance relation $\lhd^*$ (reflexive and transitive closure of immediate dominance; see Backofen et al. 1995), which allows us to describe the syntactic information common to (16a) and (16b) in (18), graphically represented as a "quasi tree" in (19):

(18) $\{N'(P_0), N'(P_1), N'(P_2), N'(P_3), N'(P_4), AP(P_5), N(P_7), PP(P_6),$
$P_0 \lhd^* P_1, P_0 \lhd^* P_2, P_1 \lhd P_5, P_1 \lhd P_3, P_2 \lhd P_4, P_2 \lhd P_6, P_3 \lhd^* P_7, P_4 \lhd^* P_7\}$



(19)

Our description language allows for a straightforward "quasi-compositional" interpretation of the quasi tree, by forming a constraint set according to the rules given in (20). We take X_i to be the semantic representation for node P_i. For the time being, we use object language constants and variables as meta-language type-0 constants (denoting themselves).

(20) (i) For every lexical node P_k, add $X_k = \alpha$, where α is the name for the object language constant representing P_k's lexical semantic value.

(ii) For each binary branching local tree, add a constraint that identifies the semantics of the mother node with the result of the function application between its daughters.

(iii) For each pair of nodes P_i and P_j for which the dominance relation is stated, we add the constraint $X_i = C(X_j)$

The interesting clause of (20) is (iii), introducing constraints which express that the semantic value of P_i is the result of the application of some (otherwise unspecified) semantic operation to the semantic value of P_j.

Application of (20) to (18)/(19) yields the constraint set (21), which can be transformed to (22) by simple substitutions.

(21) $\{X_5 = former', X_6 = in\_Amsterdam', X_7 = professor',$
$X_1 = X_5@X_3, X_2 = X_6@X_4, X_0 = C_1(X_1), X_0 = C_2(X_2), X_3 = C_3(X_7), X_4 = C_4(X_7)\}$

(22) $\{X_0 = C_1(former'@C_3(professor')),$
$X_0 = C_2(in\_Amsterdam'@C_4(professor'))\}$

(22) gives a compact underspecified description of the semantically ambiguous expression (15), saying that the semantic value of X_0 is the result of applying some (otherwise unspecified) semantic operation to $former'(\alpha)$, where α results from application of some other semantic operation to $professor'$, and at the same time the result of applying some semantic operation to $in\_Amsterdam'(\beta)$, where β results from application of some other semantic operation to $professor'$. By appropriate substitution of the function variables, it can be verified that the two readings (17a) and (17b) are in fact covered by the description. Application of (23a) to (22) leads to (17a) as value for X_0, whereas by applying (23b) we obtain (17b).

(23) (a) $\{C_1 \leftarrow \lambda X.X, C_2 \leftarrow \lambda X.former'@X,\ C_3 \leftarrow \lambda X.in\_Amsterdam'@X,$
$C_4 \leftarrow \lambda X.X\}$

(b) $\{C_1 \leftarrow \lambda X.in\_Amsterdam'@X,\ C_2 \leftarrow \lambda X.X,\ C_3 \leftarrow \lambda X.X,$
$C_4 \leftarrow \lambda X.former'@X\}$

3 Additional Restrictions on the Formalism

Straightforward and intuitively appealing as the treatment of the example is, it is far from guaranteeing adequate results. The constraint set (22) covers the two wanted solutions of our example. However, as brief consideration shows, (22) covers a wide variety of undesirable readings at the same time. There are several different sources for inadequate results.

(24) $\{C_1 \leftarrow \lambda X.\mathit{it\_rains}', C_2 \leftarrow \lambda X.\mathit{it\_rains}'\}$

First, the substitution (24) renders $X_0 = \mathit{it\_rains}'$ as a solution for (22). Given our formalism, as it has been defined so far, the solution goes through, although it is completely unrelated to the semantic information which is encoded in the constraint set. This means that we have to take care that operator variables be instantiated only in a way that preserves the information contained in the arguments. We will come back to the requirement in a minute, and see that it has wide-reaching consequences. In advance, we would like to briefly discuss two other problems that arise for our approach.

One of the problems is that we have to take object-language type compatibility into account as an additional source for restrictions, because unconstrained substitution may lead to solutions that describe ill-formed object-language expressions.

(25) $\{X_0 = X@Y, X_0 = C_1(\mathit{walk}'), X_0 = C_2(j^*)\}$

(25) should contain sufficient information to determine $\mathit{walk}'@j^*$ as the unique meaning representation, given closure information. However, we get the ill-formed $j^*@\mathit{walk}'$ as well. We will not provide an explicit treatment for OL type information, here, but rather take it that OL type information can be integrated as sortal information in the description language. We assume that OL type constraints will be reflected in terms of sort descriptions in USDL in roughly the following way: Type-0 constants are assigned their OL type; type-0 variables are in general assigned Top, but type information contained in NL syntactic categories may be reflected in sort restrictions for the associated USDL variables; a USDL function constant like the two-place @, e.g., is assigned $(<\sigma,\tau>,\sigma) \rightarrow \tau$ as sort description, where σ, τ are variables ranging over sorts. OL type clashes do not affect syntactic well-formedness of DL expressions, but lead to the assignment of a designated semantic value (of sort Bottom). We hope that computation of the kind of sort information we need can be integrated into our constraint solving mechanism using standard first-order unification. We have not looked into the topic in greater detail, however.

There is a further problem which, however, is in fact a minor one. Instead of the substitutions in (23), we could have chosen substitutions that introduce function variables into any position in the expression. The most general readings indeed seem to be (27a) and (27b), which are generated by (26a) and (26b), respectively.

(26) (a) $\{C_2 \leftarrow \lambda X.C_1(\mathit{former}'@C(X)),$
$C_3 \leftarrow \lambda X.C(\mathit{in\_Amsterdam}'@C_4(X))\}$
(b) $\{C_1 \leftarrow \lambda X.C_2(\mathit{in\_Amsterdam}'@C(X),$
$C_4 \leftarrow \lambda X.C(\mathit{former}'@C_3(X))\}$

(27) (a) $X_0 = C_1(\mathit{former}'@C(\mathit{in\_Amsterdam}'@C_4(\mathit{professor}')))$
(b) $X_0 = C_2(\mathit{in\_Amsterdam}'@C(\mathit{former}'@C_3(\mathit{professor}')))$

This inadequacy is not due to the formalism itself, but results from the way in which the input information has been presented. The syntactic constraint set (18), which served as the basis for the semantic constraints in (21) and (22), respectively, contains positive statements about syntactic relations only. Thus, it does not exclude the existence of additional intervening linguistic material. Accordingly, the semantic constraints tell us only about the relative scope ordering of the OL constants *former'*, *in_Amsterdam'*, and *professor'*, and do not exclude additional semantic objects to appear at any level of the utterance representation. If we assume

that the linguistic material is completely covered by the syntactic constraints, we are entitled to perform a closure operation, i.e. instantiate each function variable in the solutions with $\lambda X.X$, which, in the case of our example, takes us to the wanted variable-free representations in (17).

We come back to the first problem, which was illustrated in (24). The moral of the example is that no vacuous abstraction, as in $\lambda X.\mathit{it\_rains}$, should be allowed: NL semantic operations can do a variety of different things to their arguments, but they do not delete them or overwrite them with different material.

It also seems to be desirable to exclude uncontrolled proliferation of semantic objects due to copying. Take for example constraint set (28), which corresponds to an information state where the word forms *John* and *likes* have been recognized, and like has been identified as main verb of the sentence.

(28) $\{X_0 = \mathit{like}'@Y@X, X_0 = C(j^*)\}$

cases $X_0 = \mathit{like}'@C(j^*)@X$ and $X_0 = \mathit{like}'@Y@C(j^*)$ are fine, as well as the special a solution gained by substituting C with $\lambda X.\mathit{like}@X@X$. This seems to indicate a general need to exclude multiple occurrences of bound variables in function terms, and a justification for a general linearity constraint. However, the situation is somewhat less clear than in the deletion case, for some kind of copying seems to be required in NL semantics, to model, e.g., coordination and ellipsis constructions: The VP *sing and dance* should have a representation like $\lambda x.\mathit{sing}'(x) \wedge \mathit{dance}'(x)$.

Our policy of distinguishing between the levels of object language and description language suggests the following view: OL abstraction and USDL abstraction are completely different notions. In fact, the USDL representation of $\lambda x.\mathit{sing}'(x) \wedge \mathit{dance}'(x)$ should be something like $L_x(\mathit{Conj}(\mathit{sing}'@x, \mathit{dance}'@x))$, where L_x and Conj are USDL function constants referring to the OL operations of abstraction with variable x and conjunction, respectively. Object-language λ-abstraction can take any form that may be required by NL meaning structure. On the other hand, USDL λ-terms are required to be linear throughout: Each variable abstracted over occurs exactly once in the body of the λ-expression. It is an interesting question how much of OL λ-conversion can be integrated into the meta-language equality concept. This question is, however, beyond the scope of this paper.

Linearity makes the number of occurrences of semantic objects a significant feature of the constraint set, since it guarantees that description language β-reduction never involves a change in the number of occurrences of a certain expression. Therefore, the difference between multiple reference to the same occurrence and reference to multiple occurrences of an OL expression becomes significant. To make the difference expressible, we take type-0 meta-language constants to be names for occurrences of OL symbols, rather than for OL symbols themselves. Thus, we will in principle refer to different occurrences of the OL (constant or variable) symbol a by $\underline{a}_i$ and $\underline{a}_j$, where $i, j \in N$ and $i \neq j$. Correspondingly, we discriminate different applications of semantic operations, and reinterpret first-order variables as ranging over occurrences.

Thus, (29a) and (29b) are significantly different constraint sets, which specify different sets of OL expressions.

(29) (a) $\{X = \underline{\mathit{like}'}_1@_1Y@_2X, X_0 = C_1(\underline{j^*}_1), X_0 = C_2(\underline{j^*}_1)\}$
(b) $\{X_0 = \underline{\mathit{like}'}_1@_1Y@_2X, X_0 = C_1(\underline{j^*}_1), X_0 = C_2(\underline{j^*}_2)\}$

The OL formula $\mathit{like}'(j^*)(b^*)$ matches the constraints in (29a). It is incompatible with those in (29b), which denote the singleton set $\{\mathit{like}'(j^*)(j^*)\}$, given closure information. Note that $\mathit{like}'(j^*)(j^*)$ is also compatible with (29a). However, it should not be by virtue of the solution $X_0 = \mathit{like}'@j_1^*@j_1^*$ for (29a) (if possible, we omit discriminating subscripts and underlines), but by free instantiation of either X and Y with some j_i^*, $i \neq 1$. For if we take USDL constants to be names of OL occurrences, $\mathit{like}'@j_1^*@j_1^*$ does not have a consistent interpretation any more: liter-

ally understood, the term denotes an OL expression where both argument places of *like'* are filled by the same occurrence of the term j^*. Thus it seems to be a reasonable general constraint to allow only single occurrences of (individual or function) constants in USDL terms. The same holds for free first-order variables, since any possible instantiation would lead to multiple occurrences of constants. For second-order USDL variables, the situation is slightly different, since they are consistently interpretable as projections like $\lambda X.X$. Thus, we assume that in USDL a *Single-Occurrence Constraint* holds for terms with respect to any description-language constants and first-order variables (we will become more specific about the constraint later).

4 Linear Unification

Given the Linearity Constraint, the Single-Occurrence Constraint, and OL type sensitivity, our approach has a much better chance to meet the intuitive requirements. However, nothing definite can be said as long as we do not have a general method to find out what our USDL expressions precisely denote. The standard way to provide such a method is the specification of a denotational semantics. Unfortunately, this standard technique is only of limited use in our case. A real semantics that maps underspecified representations directly into some appropriate denotation based on truth-conditions for OL expressions is not only out of the range of this paper, but out of the range of current research in general, as far as structural underspecification of Levels 2 and 3 is concerned. A more modest semantics using term algebras looks good at first sight, since it assigns USDL expressions sets of OL terms as denotations, and moreover could serve as the first step of an indirect, two-step process of assigning real denotations. The problem with this approach is that the denotational interpretation does not take us really closer to the actual determination of the class of described OL terms. It merely transforms the problem of identifying appropriate substitutions to the structurally similar problem of identifying appropriate assignments. We will not pursue the question of a denotational semantics in this paper, but instead try to specify a constraint-solving mechanism, which generates the possible solutions for constraint sets in a systematic way, and thus serves as an operational semantics for USDL.

There is a constraint-solving algorithm for higher-order constraints, i.e., the higher-order unification (HOU) algorithm of Huet 1975. However, this does not mean that our task has already been solved. On the one hand, unconstrained HOU is an extremely unhandy tool to work with: not only is it semi-decidable, but also has a strong tendency towards combinatorial explosion. On the other hand, we have to make sure that the algorithm works under the linearity constraint for USDL. In this section, we will first give a short outline of the basic mechanism of Huet's algorithm, and then try to adapt it to the needs of our description language. The resulting formalism of "Linear Unification" will also exhibit a much better computational behaviour.

4.1 Huet's Algorithm

The general problem of HOU corresponds to the standard FO unification problem: The global goal is to solve a set of equations. The basic subtask is to find substitutions which render both sides of an equation $\alpha = \beta$ identical. The general form for the equations is given by $\alpha = \lambda X_1 ... X_k.\Psi(\alpha_1, ..., \alpha_n)$ and $\beta = \lambda X_1 ... X_k.\Phi(\beta_1, ..., \beta_m)$. The important difference to first-order unification is that the head symbols of Ψ and Φ need not be constants, but can also be bound or free variables. Bound variables as head symbols behave like constants; terms with constant or bound-variable head symbol are called *rigid*. *Flexible* terms with free variables as head symbols are the

really novel feature of HOU. For the sake of simplicity, we consider only cases of terms without λ-prefix (and therefore without the possibility of bound variables), as in (30). The extension to the general case is straightforward.

(30) (a) $\alpha = \Psi(\alpha_1, ..., \alpha_n)$
(b) $\beta = \Phi(\beta_1, ..., \beta_m)$

We have to distinguish three sub-cases:

(i) α and β are rigid (Ψ , Φ constants)
(ii) α is flexible, β is rigid (Ψ variable, Φ constant)
(iii) α and β are flexible (Ψ , Φ variables)

Case (i) corresponds to conventional first-order unification with constant functor symbols k_1, k_2. If $k_1 \neq k_2$, unification fails. If $k_1 = k_2$, the unification task is reduced to the set of smaller tasks $\alpha_i = \beta_i (1 \leq i \leq m = n)$ (*Decomposition*). E.g., $\{walk'@X = Y@j^*\}$ is decomposed to $\{walk' = Y, X = j^*\}$. As in the first-order case, decomposition is deterministic.

Case (ii): The general idea for processing flexible-rigid combinations is to reduce them to the decomposition case by appropriate instantiation of the head variable . The task is to bring the head constant k of the rigid term to the head position of the flexible term α, which can be done in essentially two ways: first, by *Projection*, i.e., by instantiating the head variable C of the flexible term to a function that projects one of its arguments; second, by *Imitation*, i.e.,by instantiating C to a λ-term with head symbol k. Take the equation $C(walk') = walk'$ as an example. A projection solution is $C = \lambda X.X$, an imitation solution $C = \lambda X.walk'$.

This simple example already shows that some non-determinism comes into play. In certain cases the Projection rule and the Imitation rule can be alternatively applied. Moreover, if the head variable is many-place and argument head symbols are variables, there are different ways to apply the projection rule. The most important source for non-determinism, however, is that the rules of projection and imitation are specified in a very general way (and must be so specified in order to make the algorithm complete): In the imitation case, e.g., variable C of our example cannot be instantiated just to the constant k. It may well be that C and k are of different type (which means in our second-order language that they are of different arity, as in $C(walk') = walk'@j^*$), and the arguments of C may contribute to the arguments of k in any grouping and order. To cover all possible cases, the imitation rule requires C to be instantiated with the *most general τ-term with head symbol k* (where τ is the type of C). If C is of type $< \sigma_1, ..., \sigma_n, \sigma >$, k of type $< \tau_1, ..., \tau_m, \sigma >$, this is $\lambda\overrightarrow{X}.k(C_1(\overrightarrow{X}), ..., C_m(\overrightarrow{X}))$, where $\overrightarrow{X} = X_1, ..., X_n$, the X_i are variables of type σ_i, and the C_j are new variables of type $< \sigma_1, ..., \sigma_n, \tau_j >$.

In (31), a simple example is given to illustrate the working of the Imitation rule.

(31)
$C(walk', j^*) = walk'@j^*$
$\Rightarrow C = \lambda XY.C_1(X,Y)@C_2(X,Y)$ (Imitation)
$\Rightarrow C_1(walk', j^*)@C_2(walk', j^*) = walk'@j^*$ (Binding)
$\Rightarrow C_1(walk', j^*) = walk', C_2(walk', j^*) = j^*$ (Decomposition)
$\Rightarrow C_1 = \lambda XY.X, C_2 = \lambda XY.Y$ (Projection)
$\Rightarrow C = \lambda XY.X@Y$ (Binding)

The imitation step itself is deterministic. However, it gives rise to quite a bit of non-determinism already in this trivial example: In the first step, the arguments of C are copied m-times (where m is the arity of k), in the second step arbitrarily many of the occurrences can be removed by an appropriate specification of the C_i functions. This also leads to many solutions which are unwanted in the view of our intended application.

Case (iii): The idea for processing the equation of two flexible terms is to reduce it to a case (ii) problem by substituting the left head variable with the most

general τ-term with head symbol k, for some arbitrarily chosen constant k of the language (*Explosion*). This amounts to a recursive Generate-and-test procedure. The effect on the computational properties needs no comment.

In the following sub-sections, we will discuss the necessary changes that the HOU algorithm must undergo, in order to meet the requirements of a constraint solving algorithm for USDL. We will present proposals for a "Linear Imitation" rule, as well as a rule of "Flexible Decomposition" which is supposed to replace Explosion.

The Projection rule does not require a special treatment. Since USDL is second-order, only projection terms with first-order head variables are possible, i.e., terms of the form $\lambda X_1 ... X_n . X_i$. By the linearity constraint, the possible projection terms further reduce to the trivial case $\lambda X.X$.

4.2 Linear Imitation

The linearity constraint on function terms requires a modification of the imitation rule, for every more than one-place head constant on the right-hand side of the equation induces λ-terms with multiple occurrences of bound variables. Thus adding the linearity constraint would reduce the application of Imitation to one-place cases. We call our modified Imitation rule "Linear Imitation"[3]. The head variable C of the flexible term is not instantiated with the most general τ-term with head symbol k of the rigid term, but rather with a member of the set of most general *linear* τ-terms with head symbol k, which is for $\alpha = C(\alpha_1, ..., \alpha_n), \beta = k(\beta_1, ..., \beta_m)$ some $\lambda \overrightarrow{X} . k(C_1(\overrightarrow{X}_1), ..., C_m(\overrightarrow{X}_m))$, such that $\overrightarrow{X_1} ... \overrightarrow{X_m} \in perm(\overrightarrow{X})$. We give the Linear Imitation ("LI") version of Example (31) in (32) below.

(32)
$$
\begin{array}{ll}
C(walk', j^*) = walk'@j^* & \\
\Rightarrow C = \lambda XY.C_1(X)@C_2(Y) & \text{(LI)} \\
\Rightarrow C_1(walk')@C_2(j^*) = walk'@j^* & \text{(Binding)} \\
\Rightarrow C_1(walk) = walk', C_2(j^*) = j^* & \text{(Decomposition)} \\
\Rightarrow C_1 = \lambda X.X, C_2 = \lambda X.X & \text{(Projection)} \\
\Rightarrow C = \lambda XY.X@Y & \text{(Binding)}
\end{array}
$$

Linear Imitation is clearly non-deterministic, since in general there is no unique most general linear τ-term. In our example, there are several possible ways to apply the LI rule, according to the fact that there is no unique most general term, but rather four of them, i.e., $\lambda XY.C_1(X)@C_2(Y)$, $\lambda XY.C_1(Y)@C_2(X)$, $\lambda XY.C_1(X,Y)@Z$, and $\lambda XY.Z@C_2(X,Y)$. Globally, however, non-determinism is reduced, since all subsequent choices of alternative projections are removed. In fact, it is the linear imitation rule that renders non-trivial projection dispensible, since it anticipates the different projection alternatives that would have been required by general imitation, and coordinates them by sorting out those cases which are incompatible with the linearity constraint.

4.3 Flexible Decomposition

The part of the general HOU algorithm which has the best chance to lead to a computational desaster is the Explosion rule. If in $C_1(\alpha_1, ..., \alpha_n) = C_2(\beta_1, ..., \beta_m)$ both C_1 and C_2 are variables, the algorithm recursively produces all kinds of compatible instantiations of C_1 and C_2, respectively. It must do so in order to systematically generate those instantiations which satisfy all the constraints. The linearity constraint requires us to modify the Explosion rule in a way that only linear substitutions for C_1 are guessed. This modification guarantees correctness with respect to the linearity constraint, but it does not lead to an essential improvement of computational behaviour.

3. The idea is due to Michael Kohlhase (personal communication).

However, linearity allows a more radical modification of the Explosion rule. Look at the simple case $X_0 = C_1(\alpha) = C_2(\beta)$, which is also the most typical one in our intended applications. Given linearity of the function terms, there are precisely three ways in which α and β can be related.

First, α and β may form disjoint parts of the resulting term. In this case, the solution should be a two-place constraint $X_0 = H(\alpha, \beta)$, which just collects the semantic information contained in $C_1(\alpha)$ and $C_2(\beta)$ under a common variable, and in addition expresses the information that α and β form disjoint parts of the described semantic structure.

Second, α and β may share material, or "overlap". Then, either α is completely contained in β, or β is completely contained in α. In the former case, β is the result of applying some operation to $\alpha : \beta = C'(\alpha)$, and C_1 can be analysed as composition of C_2 and some new function variable C' ($C_1 = \lambda X.C_2(C'(X))$). The latter case works correspondingly: We obtain $\alpha = C'(\beta)$ and $C_2 = \lambda X.C_1(C'(X))$.

This solution cannot be straightforwardly generalized to flexible many-place terms, for the three cases, disjointness and inclusion in either direction, may simultaneously apply. Also, the two head variables do not necessarily stand in the simple relation that one completely includes the other. In $C_1(\alpha, \beta) = C_2(\gamma, \delta)$, it may well be the case that α includes γ, where β is included by δ ($\alpha = C'(\gamma), \delta = C''(\beta)$). This requires C_1 and C_2 to be properly overlapping functions with a common head part. In (33), we give a general formulation of what we call the rule of Flexible Decomposition (FD), which is able to alternatively generate the full range of disjoint, overlapping, and mixed cases, in one step.

(33) Flexible Decomposition:

Replace $C_1(\alpha_1, ..., \alpha_n) = C_2(\beta_1, ..., \beta_m)$ by

(i) equations A_i for every i: $1 \leq i \leq k$ (for some $k \leq m + n$), which are either of the form $C(\beta_{j1}, ..., \beta_{jl}) = \alpha_j$, for some $l \geq 0$, or of the form $C(\alpha_{j1}, ..., \alpha_{jl}) = \beta_j$,for some $l \geq 0$, such that the C are pairwise different new variables of arity ≥ 0, all argument terms of C_1 and C_2 occur exactly once in one of the A_i, and

(ii) the equations $C_1 = \lambda X_1 ... X_n . H(\gamma_1, ..., \gamma_k)$ and $C_2 = \lambda Y_1 ... Y_m . H(\delta_1, ..., \delta_k)$, where H is a new k-place function variable, $\sigma(\gamma_i)$ is the α side, and $\sigma(\delta_i)$ is the β side of A_i, for $1 \leq i \leq k$, $\sigma = \{\alpha_l \leftarrow X_l | 1 \leq l \leq n\} \cup \{\beta_l \leftarrow Y_l | 1 \leq l \leq m\}$

Note that the choice of the A_i completely determines the equations for C_1 and C_2.

First, let us look at an example which illustrates the working of the rule. For equation (35), we give two possible solutions, where (35) is the maximally disjoint one, and (36) is an example for a solution that realizes all possible relations between the arguments of C_1 and C_2. The (a) clauses of (35) and (36) show the immediate output of the FD rule. The (b) clauses show the intermediate result obtained for X_0 by Binding.

(34) $X_0 = C_1(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = C_2(\beta_1, \beta_2, \beta_3)$

(35) (a) $Z_1 = \alpha_1, Z_2 = \alpha_2, Z_3 = \alpha_3, Z_4 = \alpha_4, U_1 = \beta_1, U_2 = \beta_2, U_3 = \beta_3,$
$C_1 = \lambda X_1 X_2 X_3 X_4 . H(X_1, X_2, X_3, X_4, U_1, U_2, U_3)$
$C = \lambda Y_1 Y_2 Y_3 . H(Z_1, Z_2, Z_3, Z_4, Y_1, Y_2, Y_3)$

(b) $X_0 = H(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_1, \beta_2, \beta_3)$

(36) (a) $Z = \alpha_1, C'(\alpha_2, \alpha_3) = \beta_1, C''(\beta_2) = \alpha_4, U = \beta_3,$
$C_1 = \lambda X_1 X_2 X_3 X_4 . H(X_1, C'(X_2, X_3), X_4, U)$
$C_2 = \lambda Y_1 Y_2 Y_3 . H(Z, Y_1, C''(Y_2), Y_3)$

(b1) $X_0 = H(\alpha_1, C'(\alpha_2, \alpha_3), \alpha_4, \beta_3)$

(b2) $X_0 = H(\alpha_1, \beta_1, C''(\beta_2), \beta_3)$

In (37a) to (37c), we show the three alternative solutions obtained by application of Rule (33) to the $X_0 = C_1(\alpha) = C_2(\beta)$ case, which we discussed above.

(37) (a) $Z = \alpha, Y = \beta, C_1 = \lambda X.H(X,Y), C_2 = \lambda Y.H(X,Y)$
(b) $C'(\alpha) = \beta, C_1 = \lambda X.H(C'(X)), C_2 = \lambda X.H(X)$
(c) $C'(\beta) = \alpha, C_1 = \lambda X.H(X), C_2 = \lambda X.H(C'(X))$

We can eliminate the use of the new variable H in the overlap clauses (37b) and (37c) because of $C_2 = \lambda X.H(X)$ and $C_1 = \lambda X.H(X)$, respectively. We thus arrive at $\{C_1 = \lambda X.C_2(C'(X)), C'(\alpha) = \beta\}$ and $\{C_2 = \lambda X.C_1(C'(X)), C'(\beta) = \alpha\}$, respectively, and will take these as the direct output alternatives of the FD rule for the special case of two one-place terms, in order to keep the notation simple.

4.4 Computational Properties of Linear Unification

The modified rules presented in the last two sections allow us to accommodate HOU to the needs of USDL. Accommodation is possible in a more cautious and a more radical way. The cautious variant consists in just replacing Imitation by Linear Imitation and Explosion by its linearized variant. The resulting system is complete. It is also correct, if we disregard the Single-Occurrence Constraint. The latter can be taken care of by a filter on the output, which sorts out all solutions containing terms with multiple occurrences of constants and first-order variables.

This seems to be an appropriate place to give the Single-Occurrence Constraint a precise definition: A USDL term must not contain more than one occurrence of any constant or variable, except one-place function variables. Multiple occurrences of one-place function variables do not necessarily lead to the duplication of occurrences, for they can be substituted by $\lambda X.X$. Since this is also the only way to make sense of them, we are entitled to add a rule that carries out the substitution as soon as a function variable occurs twice, and thus eliminates both occurrences.

The problem with the cautious variant of the algorithm is, however, that it does not terminate, and moreover exhibits a similar tendency to combinatorial explosion, as the original unconstrained version does. From a computational point of view, the radical variant is much more attractive. It uses Linear Imitation, and Flexible Decomposition in the place of Explosion. Moreover, the Single Occurrence Constraint can be used as an on-line filter, which immediately blocks a path in the search tree, as soon as a term is generated which violates the constraint. We expect this version of the algorithm to terminate, although the corresponding proof has not yet been carried out.

We also have not proved completeness under the additional restrictions, but are confident about it. At least we can say that FD plus Single-Occurrence Constraint lead to the elimination of several notorious sources of undesirable computational behaviour. In the beginning of the last section, we have already mentioned the possibly infinite recursive generation of possible instantiations as a motivation for the FD rule. Linear Imitation may also lead to infinite recursion in the case of independent multiple occurrences of a function variable in the constraint set. Look,e.g., at $F(g(a)) = g(F(a))$. Here immediate projection ($F = \lambda X.X$), which unifies the two occurrences of g, is not the only solution. LI renders $F = \lambda X.g(C(X))$, which leads to $g(C(g(a))) = g(g(C(a)))$, and further to $g(g(a))$, as a solution, as well as to $g^n(a)$ for any natural number n, by iteration of the LI rule. The Single-Occurrence Constraint filter excludes all but the first alternative, in accordance with the basic intuitions discussed in Section 3, and at the same time blocks the recursion process in its initial step.

Finally, we would like to point to the interaction between FD and Single-Occurrence Constraint. FD reduces the substitution of the head variable in a pair of flexible terms to a finite number of alternatives - a number, however, that drastically increases with the arity of the head variables. But many-place function variables are only introduced to cover the disjointness option for the arguments. Thus, the constraint automatically blocks the computationally expensive introduction of

function variables with higher arity, if the arguments of the flexible terms overlap—and many of the typical examples we want to treat, syntactic ambiguities as well as scope underspecification, are such overlap cases.

5 Examples

In this section, we will demonstrate the application of the linear unification formalism to our syntactic underspecification example, as well as to cases of scope underspecification.

5.1 Syntactic Underspecification

We start with the *former professor in Amsterdam* example. We will be as explicit as possible, for this first example, indicating not only the changes certain formulae in the constraint set undergo, but also keeping track of the constraint set as a whole. For notational convenience, we introduce a numbering for the different occurrences of formulae (written in square brackets).

We start with constraint set (38), where we assume that the two description language occurrences of $prof'$ refer to the same occurrence of the OL constant $prof'$, and that the two occurrences of @ denote different instances of functional application, without indexing them explicitly. (38) is transformed to (39) by binding ([2'] being the result of the substitution of X_0 in [2] by $C_1(former'@C_3(prof'))$, the right-hand side of equation [1].

(38) [1] $X_0 = C_1(former'@C_3(prof'))$, [2] $X_0 = C_2(in\_A'@C_4(prof'))$
(39) [1],[2'] $C_1(former'@C_3(prof')) = C_2(in\_A'@C_4(prof'))$

Application of the Flexible Decomposition rule (special variant) to [2'] in (39) yields (40) with the two new constraints [3] and [4], and binding plus β-reduction transforms [1] to [1'] of (41).

(40) [1],[3] $C'(former'@C_3(prof')) = in\_A'@C_4(prof')$,
[4] $C_1 = \lambda X.C_2(C'(X))$
(41) [1'] $X_0 = C_2(C'(former'@C_3(prof')))$, [3], [4]

By choosing the other overlap variant of FD, we would have obtained (42) and (43) instead, which lead to the second solution for our constraint set.

(42) [1], [3] $C'(in\_A'@C_4(prof')) = former'@C_3(prof')$,
[4] $C_2 = \lambda X.C_1(C'(X))$
(43) [1'] $X_0 = C_1(C'(in\_A'@C_4(prof')))$,[3],[4]

Since the computation of the second alternative is completely analogous to that of the first reading, we do not present it explicitly here. Application of the third FD option to (39) is blocked, since right side and left side of [2'] overlap (because of shared occurrence of the constant $prof'$).

There are two alternative ways to apply the Linear Imitation rule to [3] of (40)/(41), since a one-place predicate is to be matched with a two-place constant. The first one results in adding constraint [5] $C' = \lambda X.Z@C''(X)$ to (41). The resulting constraint set is transformed to (44), (45), and (46), by the deterministic steps of Binding with [5] and β-reduction, (Standard) Decomposition of [3'], and another step of Binding with [6], respectively.

(44) [1''] $X_0 = C_2(Z@C''(former'@C_3(prof')))$,
[3'] $Z@C''(former'@C_3(prof')) = in\_A'@C_4(prof')$,
[4'] $C_1 = \lambda X.C_2(Z@C''(X))$, [5] $C' = \lambda X.Z@C''(X)$
(45) [1''],[4'],[5],[6] $Z = in\_A'$, [7]$C''(former'@C_3(prof')) = C_4(prof')$
(46) [1'''] $X_0 = C_2(in\_A'@C''(former'@C_3(prof')))$,
[4''] $C_1 = \lambda X.C_2(in\_A'@C''(X))$, [5'] $C' = \lambda X.in\_A'@C''(X)$, [6], [7]

The other LI alternative would have been $C' = \lambda X.C''(X)@Z$, instead of [5]. As you can easily check, processing along this branch stops after another two steps.

since decomposition renders equation $C''(former'@C_3(prof')) = in\_A'$, which has no linear solution. Special FD, which transforms [7] to [8] and [9], makes (46) into (47).

(47) [1‴], [4″], [5′], [6],
[8] $C'''(prof') = former'@C_3(prof')$, [9]$C_4 = \lambda X.C''(C'''(X))$

Note that the other possible result of applying special FD to [7] would lead to the unsolvable equation $C'''(former'@C_3(prof')) = prof'$. - Next, we apply LI to [8], which adds [10] $C''' = \lambda X.U@C''''(X)$, bind with [10], which transforms [8] to [8′] $U@C''''(prof') = former'@C_3(prof')$, decompose [8′] into [11] and [12], and bind with [11], which renders (48).

(48) [1‴], [4″], [5′], [6], [9′] $C_4 = \lambda X.C''(former'@C''''(X))$,
[10′]$C''' = \lambda X.former'@C''''(X)$, [11] $U = former'$, [12] $C''''(prof') = C_3(prof')$

Application of special FD to [12] yields [13] $C''''(prof') = prof'$ and [14] $C_3 = \lambda X.C''''(C''''(X))$ (note that in this case both FD alternatives render identical results); application of Projection to [13] adds [15] $C'''' = \lambda X.X$, and by binding with [14] and [15] we obtain (49).

(49) [1⁗] $X_0 = C_2(in\_A'@C''(former'@C''''(prof'))))$, [4″], [5′], [6], [9′], [10′],
[11], [13′] $prof' = prof'$, [14], [15]

[13′] as a tautology can be removed from the constraint set, which completes the derivation. The other branch originating from (42) and (43) would have led to a solution containing (50).

(50) $X_0 = C_1(former'@C''(in\_A'@C''''(prof')))$

By the closure operation argued for in Section 5—substituting all function variables with $\lambda X.X$, we obtain the readings of former professor in Amsterdam, which are repeated here as (51).

(51) (a) $former'@(in\_Amsterdam'@professor')$
(b) $in\_Amsterdam'@(former'@professor')$

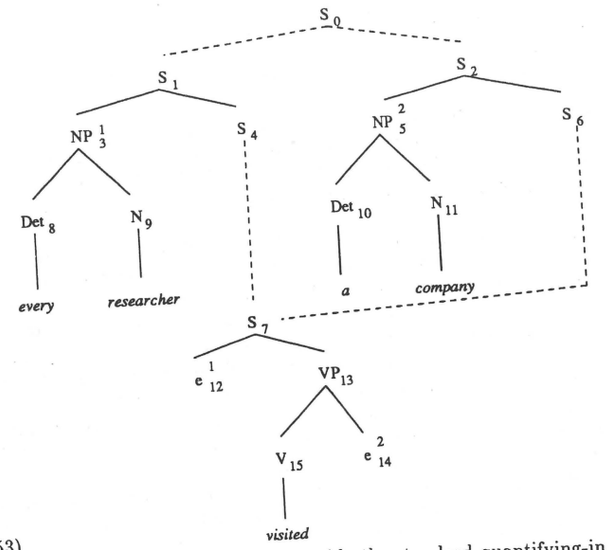
As you may check, these are in fact the only readings which can be obtained from the original constraint set.

5.2 Scope Underspecification

The example of the last sub-section in fact was a case of scope underspecification, already. Two alternative scope orderings of the two modifiers correspond to the different syntactic analyses in terms of constituent structure, which can be assigned to the N' phrase in question. Insofar, as we analyse scope ambiguities of quantifiers and other sentential operators as being based on syntactic ambiguities, which is the line of Montague 1973 and May 1985, we can just carry over our analysis. In (53), we give the quasi-syntactic tree representing the two scope readings of Sentence (52)[4].

(52) *Every researcher visited a company*

4. We do not care about linear precedence information here. —The syntactic representation is taken over from Muskens 1995.



(53)

A slight extension is necessary to provide the standard quantifying-in treatment. We add the two clauses of (54) to the interpretation principles given in (20). The symbols x_i and x_j are USDL constants that stand for the corresponding OL first-order variables x_i and x_j, respectively, not for two different occurrences of x. L_{xj} is taken as a one-place USDL function constant denoting the OL abstraction operation with x_j.

(54) (i) For each terminal node P_k which is occupied by an indexed empty element e^i add the constraint $X_k = x_i$.
(ii) For each binary branching node P_k with daughter nodes P_l and P_m, where P_l is an NP with index j, add the constraint $X_k = X_l@L_{xj}(X_m)$

To avoid notational complexity, we take *every_researcher'* to be short for the USDL term which uniquely specifies the OL expression $\lambda F[\forall x(researcher'(x) \rightarrow F(x))]$. Likewise, we use *a_company'* and *visit'* as shortcut names for $\lambda G[\exists y(company'(y) \rightarrow G(y)]$ and $\lambda u\lambda v[visit'(v)(u)]$.

We assume that the computation of the internal semantics of the noun phrases and the constituent S_7 has been carried out already, and thus start with a subset of the syntactic constraint set, which is graphically represented in (55):

(55) $\{S_0 \lhd^* S_1, S_0 \lhd^* S_2, S_4 \lhd^* S_7, S_6 \lhd^* S_7, S_1 \lhd S_3, S_1 \lhd S_4, S_2 \lhd S_5, S_2 \lhd S_6\}$

The semantic constraint set corresponding to (55) is the union of the sets specified in (a), (b), and (c) of (56), where (56a) contains the pre-computed parts of the description, (56b) the constraints gained by application of (54)(ii), and (56c) the constraints induced by the $\lhd^*$ clauses of (55) according to (20)(iii).

(56) (a) $X_3 = every\_researcher', X_5 = a\_company', X_7 = visit'@x_1@x_2$
(b) $X_1 = X_3@L_{x1}(X_4), X_2 = X_5@L_{x2}(X_6)$
(c) $X_0 = C_1(X_1), X_0 = C_2(X_2), X_4 = C_3(X_7), X_6 = C_4(X_7)$

From these constraints we easily arrive at (57).

(57) (a) $X_0 = C_1(every\_researcher'@L_{x1}(C_3(visit'@x_1@x_2)))$
(b) $X_0 = C_2(a\_company'@L_{x2}(C_4(visit'@x_1@x_2)))$

At this stage, the constraint set has essentially the same structure as the corresponding set (38) for the example discussed in Section 5.1. Accordingly, the algorithm works in precisely the same way, resulting in the pair of solutions given

in (58), which are transformed by closure to (59).

(58) $X_0 =$
$C_1(every\_researcher'@L_{x1}(C'(a\_company'@L_{x2}(C''(visit'@x_1@x_2)))))$
$X_0 =$
$C_2(a\_company'@L_{x2}(C'(every\_researcher'@L_{x1}(C''(visit'@x_1@x_2)))))$

(59) $X_0 = every\_researcher'@L_{x1}(a\_company'@L_{x2}(visit'@x_1@x_2))$
$X_0 = a\_company'@L_{x2}(every\_researcher'@L_{x1}(visit'@x_1@x_2))$

The pair of solutions in (59) describes the two possible scope readings of sentence (52).

There are well-known arguments to the end that scope ambiguity should not be regarded as an effect of an underlying ambiguity at some abstract syntactic level, but rather as a genuinely semantic ambiguity that results from the analysis of an unambiguous close-to-surface syntactic representation. The standard technique to realize this view is Cooper storage. The QLF formalism offers a technique to generate underspecified representations from a specific syntactic basis . Muskens 1995 proposes another treatment, which makes use of a subordination relation between labels taken over from Reyle 1993. In the following, we want to show that a direct modelling of scope underspecification can be taken to be a special application of the higher-order constraint approach of USDL.

We start by adding two new interpretation rules, one for quantifiers, and the other one for one-place sentential operators like negation. As far as treatment of scope is concerned, they replace (54). However, the latter may be necessary to interpret other uses of syntactic co-indexing (e.g., in connection with long-distance phenomena). Also, we associate two semantic meta-variables X_k and X'_k with each non-lexical syntactic node P_k (the "upper" and the "lower" semantic value of P_k), and modify the functional application rule of (20)(ii) in the way that the lower value of the mother is equated with the functional application of the upper values of the daughters. We will comment on the purpose of this change below. The two new rules are given in (60); the functions **I** and **B** will be explained below.

(60) (i) For each binary branching node P_k with daughter nodes P_l and P_m, P_l an NP, add the constraints $X'_k = X_m@x_i$ and $X_{\mathbf{I}(k)} = C(X_l@L_{xi}(C'(X'_k)))$.

(ii) For each binary branching node P_k with daughter nodes P_l and P_m, P_l a one-place sentential operator, add the two constraints $X'_k = X_m$ and $X_{\mathbf{B}(k)} = C(X_l@C'(X'_k))$.

As in the Cooper Storage treatment of quantifiers, the immediate contribution of an NP to the composition process is just an individual variable; the real quantifier denotation comes into play in a different way. In contrast to Cooper storage, however, the quantifier is not stored for later use, but immediately associated with a certain syntactic node higher up in the tree, by the second constraint. (60ii) works accordingly, except for the fact that no variable need to be introduced and abstracted over. $\mathbf{I}(k)$ is meant to render the node ID corresponding to the widest scope position which a NP daughter of P_k can take. $\mathbf{B}(k)$ provides the corresponding information for operators (which may be sensitive to different syntactic constraints). In the simple examples we are discussing in this paper, it always holds that $\mathbf{I}(k)$ and $\mathbf{B}(k)$ identify the root node, so we will not make explicit use of the notation in the following.

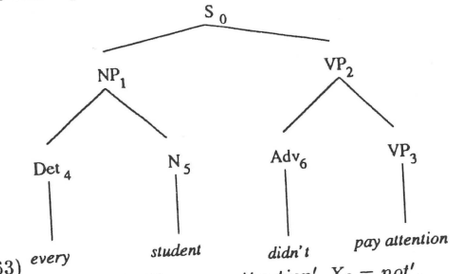
The rules in (60) in some sense correspond to the operation of quantifier storage. (61), on the other hand, is a kind of counterpart to the retrieval mechanism in Cooper Storage and related treatments of quantifiers. As mentioned above, we provide two semantic meta-variables for each syntactic constituent P_k. Clause (i) of (61) allows semantic operations to intervene between the upper value X_k and the lower value X'_k at any S node. Clause (ii) excludes this possibility for all other categories (lexical nodes are only assigned one meta-variable).

(61) (i) For every S node P_k, add $X_k = C(X'_k)$
(ii) For every other non-lexical node P_k, add $X_k = X'_k$

Note that (60) and (61) do not interact in the way that the operator must be applied at the top-most S node $P_{\mathbf{I}(k)}$ itself. S nodes lying on the path in between P_k and $P_{\mathbf{I}(k)}$ are possible candidates, as well.

As an example, we take Sentence (1) from Section 1, repeated here as (62), and assume (63) as its syntactic analysis.

(62) *Every student didn't pay attention*



(63)

(64) (a) $X'_1 = every\_student', X'_3 = pay\_attention', X_6 = not',$
(b) $X'_2 = X_3, X_0 = C_2(X_6@C_4(X'_2)))$
(c) $X'_0 = X_2@x, X_0 = C_1(X_1@L_x(C_3(X'_0)))$
(d) $X_1 = X'_1, X_2 = X'_2, X_3 = X'_3, X_0 = C_5(X'_0)$

(64) shows the constraints which we obtain from lexical insertion or pre-computation (64a), processing of the operator-containing VP (64b) and the NP-containing S (64c), as well as from the rules relating upper and lower values (64d). By simple substitutions, we obtain the three constraints in (65), and in turn by linear unification and closure the desired readings given in (66).

(65) $X_0 = C_1(every\_student'@L_x(C_3(pay\_attention'@x)))$
$X_0 = C_2(not'@C_4(pay\_attention'))$
$X_0 = C_5(pay\_attention'@x)$

(66) $X_0 = every\_student'@L_x(not'@(pay\_attention'@x))$
$X_0 = not'@(every\_student'@L_x(pay\_attention'@x))$

6 Concluding Remarks

In the last section, we have demonstrated some applications of the USDL treatment of semantic underspecification, based on the Linear Unification concept presented in Section 4. The algorithm we have employed is terminating, as we hope, and the examples show that it allows the treatment of different cases with a quite reasonable amount of computation. Also, the general situation seems to be promising in view of the fact that the maximum arity of semantic operator constants needed in NL analysis is low, maybe only 2, that the constraints we start with employ only one-place function variables, that many-place variables come into play only when disjointness is a possible option, and that the problem classes of syntactic ambiguity and scope ambiguity are definite overlap cases.

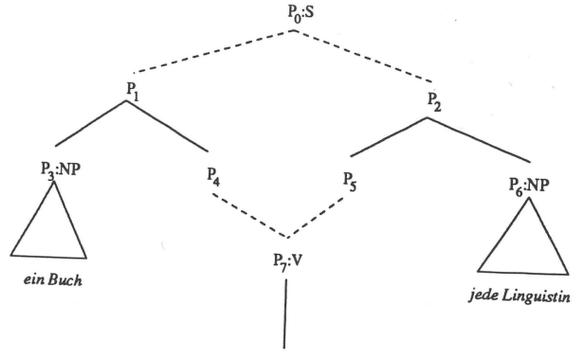
We would like to conclude the paper by sketching a few other application cases, and commenting on some open problems that become apparent in connection with the examples. First, look at the German example (67), which is a combined case of syntactic and scope underspecification.

(67) *Ein Buch zitiert jede Linguistin*

(67) has four plausible readings, namely the two pairs of scope readings that correspond to the English sentences in (68). As to the syntactic ambiguity, which

typically arises in connection with flexible German word-order, see our comments on Example (10) in Section 1.

(68) (a) *A book cites every linguist*
(b) *Every linguist cites a book*



(69)

Given (69) as underspecified syntactic representation, our interpretation rules produce a constraint set which can be shown to specify exactly the set of four plausible readings given in (70), by Linear Unification and Closure.

(70) (a) $\forall x(linguistin'(x) \rightarrow \exists y(buch'(y) \wedge zitier'(x,y)))$
(b) $\exists y(buch'(y) \wedge \forall x(linguistin'(x) \rightarrow zitier'(x,y)))$
(c) $\forall x(linguistin'(x) \rightarrow \exists y(buch'(y) \wedge zitier'(y,x)))$
(d) $\exists y(buch'(y) \wedge \forall x(linguistin'(x) \rightarrow zitier'(y,x)))$

Thus our formalism is able to treat this more complex case, too. A closer look at (69), however, shows that the treatment is not completely unproblematic. First, there is no unique assignment of syntactic categories to some node symbols of (69). If we want to exploit the type constraints on USDL variables which are induced by the syntactic category information (we do not need it in this particular example), we have to use disjunction. Second, (69) is based on the assumption of highly simplified syntactic structures, i.e. those in (70), rather than syntactically more plausible ones, like those of (72).

(71) (a) $[_S[_{NP}ein\ Buch][_{VP}zitiert[_{NP}jede\ Linguistin]]]$
(b) $[_S[_{VP}[_{NP}ein\ Buch]zitiert][_{NP}jede\ Linguistin]]$
(72) (a) $[_{S'}[_{NP}ein\ Buch]_i[_S t_i[_{VP}[_{NP}jede\ Linguistin]zitiert]]]$
(b) $[_{S'}[_{NP}ein\ Buch]_i[_S[_{NP}jede\ Linguistin][_{VP}t_i zitiert]]]$

However, it seems that a compact representation of the alternatives of (72) can only be done by making essential use of disjunction (which also holds for many other non-trivial cases of syntactic ambiguity). Accordingly, the semantic information induced by this underspecified syntactic structure can only be expressed by introducing a disjunction operator into USDL.

Finally, let us have a short look at the incompleteness type of underspecification, which is typical for spoken input. Let us consider the English variant of (13) in Section 1 as an example, which we repeat here as (73).

(73) *... logician ... cites ... paper ...*

The USDL treatment is straightforward, but does not give us a lot of insight. Given that the three fragments belong to one syntactic unit, we can establish the syntactic constraint set (74), and infer the semantic constraint set (75).

(74) $P_0 \lhd^* P_1, P_0 \lhd^* P_2, P_0 \lhd^* P_3$
(75) $X_0 = C_1(logician'), X_0 = C_2(cite'), X_0 = C_3(paper')$
(76) $X_0 = H(logician', cite', paper')$

By double application of the disjointness variant of FD, we obtain (76). (76) is the only possible solution because the other FD alternatives require overlap of the arguments, which is impossible for distinct constants. The solution is intuitively correct insofar as it encodes the amount of certain semantic knowledge we can derive on the basis of (73). However, in a case where we have a fragmentary understanding of a spoken utterance, we usually conclude considerably more about the intended semantics of the utterance, using heuristic devices like the following: "In the case of lacking counter-evidence take the recognized verb to be the main verb of the sentence, and take the recognized common nouns to be the heads of the complements of the verb". This kind of heuristics in fact has been used in AI speech understanding systems, and is known under the name of "head-driven island parsing". We can reconstruct the heuristically based inference in the following way: In absence of certain syntactic information, we guess (77) as an option for the underspecified syntactic description of (73).

(77) $[_S[_{NP}[_{Det}...][_{N'}...logician...][_{VP}cite[_{NP}[_{Det}...][_{N'}...paper...]]]]]$

The general dominance relation is indicated here by the dots surrounding *logician* and *paper*. (77) leads to the semantic description (78), as one possible scope reading, where $\mathbf{Q_1}$ and $\mathbf{Q_2}$ are Type-0 USDL variables ranging over OL determiner representations.

(78) $X_0 =$
$C_1(\mathbf{Q_1}@C_2(logician')@L_{x1}(\mathbf{Q_2}@C_3(paper')@L_{x2}(C_4(cite'@x_2@x_1))))$

By application of further heuristic rules, e.g.: "Assume an indefinite determiner in the absence of an appropriate sufficiently salient antecedent", and "Don't assume modifiers without positive evidence" (which provides the basis for default application of the closure operation), we arrive at (79) as a hypothesized reading:

(79) $X_0 = a'@logician'@L_{x1}(a'@paper'@L_{x2}(cite'@x_2@x_1))$

Two messages can be read off this example. The first one is that the higher-order approach opens an interesting perspective for a non-trivial combination of "deep" conventional and "flat" heuristic techniques in semantic analysis. It has often been stated that both kinds of processing are required to bring about robust and reliable NL understanding systems. In particular, this holds for systems which process spoken input. However, the few systems under development which combine deep and flat processing techniques at all (like the dialogue translation system Verbmobil), run them as separate modules, using different data-structures for semantic representation: If the deep analysis fails at some stage, flat analysis starts from the input string or word lattice again and tries to come up with an approximate result, disregarding any information obtained from the deep analysis. The reason is that there has been no possibility so far to provide a common framework, since logic-based, compositional interpretation seems to require full grammatical information. The approach presented here opens the way to a sensible integration of flat and deep techniques, because their operation can be modelled in a common formal framework. This allows heuristic interpretation techniques to make use of any portion of information that can be obtained from deep analysis, and vice versa; and it allows to compare the results of alternative analysis techniques in order to keep them consistent. Thus, the proposed approach may also help to bridge the gap between logically based semantics and cognitively adequate processing techniqes.

The second message conveyed by our example is somewhat less encouraging. By presenting (78) as the USDL representation of the result of using the head-driven island heuristics for parsing our example, we have cheated a little. In fact, by applying the heuristics, we arrive at a slightly more specific information state than the one expressed by (78). The assumption that *logician* and *paper* are the syntactic heads of the subject and direct object NP, respectively, implies more than just that the predicates *logician'* and *paper'* occur somewhere in the restriction of the respective quantifiers. The NP heads may occur arbitrarily deeply embedded in

function terms of the form $\alpha_1@(...@(\alpha_n@\beta)...)$, but (disregarding the possibility of relational nouns with internal arguments) only as the rightmost leaf β, since every possible intervening modifier takes them as arguments. To express this constraint, we would need a partial description of sequences of operations of arbitrary length, which is reminiscent of the "Functional Uncertainty" concept which has been developed in the context of LFG (see Kaplan and Zaenen 1989). The formalism does not allow this kind of specification so far.

We are just at the beginning of our work towards a tool for modelling radical underspecification. However, we hope that the first results are convincing enough to raise interest in higher-order constraint-based analysis of natural-language semantic information.

References

Alshawi, H. and Crouch, R.: 1992, Monotonic semantic interpretation, in *Proceedings 30th Annual Meeting of the Association for Computational Linguistics*, pp 32–38

Backofen, R., Rogers, J., and Vijay-Shanker, K.: 1995, A first-order axiomation of the theory of finite trees, *Journal of Logic, Language and Information*

Bos, J.: 1996, Predicate logic unplugged, *Proceedings of the 10th Amsterdam Colloquium*, forthcoming

Cooper, R.: 1993, Towards a general semantic framework, in R. Cooper (ed.), *DYANA-2 Deliverable R2.1.A*, pp 49–97, The DYANA-2 Project Administrator, ILLC, University of Amsterdam

Egg, M. and Lebeth, K.: 1995, Semantic underspecification and modifier, in J. Kilbury and R. Wiese (eds.), *Integrative Ansaetze in der Computerlinguistik (DGfS/CL '95), Düsseldorf, Seminar fuer Allgemeine Sprachwissenschaft*, pp 19–24, DGfS/CL

Huet, G.: 1975, An unification algoritm for typed lambda calculus, *Theoretical Computer Science*

Kaplan, R. M. and Zaenen, A.: 1989, Long-distance dependencies, constituent structure, and functional uncertainty, *Alternative Conceptions of Phrase Structure*

Kuschert, S.: 1996, A type-theoretic semantics for λ-drt, *Proceedings of the 10th Amsterdam Colloquium*, forthcoming

May, R.: 1985, *Logical Form*, MIT Press, Cambridge Mass

Montague, R.: 1973, The proper treatment of quantification in ordinary english, in J. H. e.a. (ed.), *Approaches to Natural Language*, pp 221–242, Reidel

Muskens, R.: 1993, A compositional discourse representation theory, in P. Dekker and M. Stokhof (eds.), *Proceedings of the 9th Amsterdam Colloquium*, pp 467–486, ILLC, University of Amsterdam

Muskens, R.: 1995, Order-independence and underspecification, *Dyana-2 Deliverable R2.2.C "Ellipsis, Underspecification, Events and More in Dynamic Semantics"*

Pinkal, M.: 1985, *Logik und Lexikon: Die Semantik des Unbestimmten*, de Gruyter, Berlin

Pinkal, M.: 1995, *Logic and Lexicon*, Kluwer Academic Publishers

Poesio, M.: 1994, *Semantic Ambiguity and Perceived Ambiguity*, unpublished manuscript

Reyle, U.: 1993, Vagueness and ambiguity in *drt*, *Journal of Semantics*

Skut, W.: 1995, *Underspecification-Based Representation of Noun Phrase Semantics, Diplomarbeit*, Universität des Saarlandes

van Eijck, J.: 1996, *Reasoning with Ambiguities*, Technical report, CWI Amsterdam, Ms.

Disambiguation as (Defeasible) Reasoning about Underspecified Representations

Massimo Poesio

Abstract

UNDERSPECIFIED REPRESENTATIONS have been proposed as a way of formalizing sentence disambiguation in terms of (usually, nonmonotonic) reasoning. But in fact the existing theories about the semantics of underspecified representations either fail to capture our intuitions about ambiguity, or lead one to believe that the kind of inferences used in disambiguation have nothing to do with what we usually mean by inference. As a solution, I propose a theory of underspecification based on two hypotheses: that disambiguation is the task of recovering the content of the utterance events that take place in a discourse situation; and every utterance event that took place in a conversation, including utterances of single lexical items, is recorded in the common ground. The underspecified representations I propose are representations of what took place in the discourse situation, rather than directly about the content of an utterance. It is shown that the resulting theory of underspecification offers a more satisfactory account of lexical disambiguation and is more general that existing theories, in that can also be used to account for structural disambiguation.

1 Introduction

A number of recently proposed theories of sentence interpretation make use of UNDERSPECIFIED REPRESENTATIONS: logical forms that encode multiple interpretations of an ambiguous sentence (Alshawi, 1992; Poesio, 1995; Reyle, 1993). Although languages including expressions that compactly 'pack' multiple meanings have long been used in natural language processing—at least since the LUNAR system (Woods, 1972)—this recent work is characterized by the intention of using such languages to formalize sentence disambiguation as a form of logical inference. (I am using the term 'logical' here in a more general sense that encompasses both valid and defeasible inference.) A uniform formalization of disambiguation in terms of, say, abductive reasoning would make it easy to account for the interactions among different disambiguation processes exemplified by, e.g., the interactions between scope disambiguation and reference resolution in sentences such as *If a cat sees a dog, the cat meows*; or the interactions between reference and parsing studied by Crain and Steedman (1985).

The original hope of some (e.g., (Poesio, 1991; Reyle, 1993)) was to use underspecified representations to formulate disambiguation as a monotonic inference process, in which the initial interpretation would stay underspecified

until subsequent information could be used to disambiguate.[1] Now, a look at current theories of discourse interpretation will show an almost unanymous consensus that disambiguation depends for the most part on defeasible inferences. Most theories of syntactic and lexical disambiguation these days are statistically based (Brown et al., 1991; Hindle and Rooth, 1993); and nonmonotonic logics are the foundation of current theories of speech act interpretation, such as (Perrault, 1990), or of tense interpretation, such as (Lascarides and Asher, 1991). Some theories of sentence processing that take abductive processing as their starting point include (Hobbs et al., 1990; Alshawi, 1992). Underspecified representations have been used as the foundation of such theories, most particularly at SRI Cambridge. In section §2 below I give an illustration of how disambiguation processes can be formalized in terms of defeasible reasoning over underspecified representations, according to these proposals.

But whatever a researcher's position on the role of defeasible inference on disambiguation is, (s)he will have to address the following question: what kind of inferences are those involved in disambiguation? Reasoning is normally taken to be a process that leads from true premisses to true conclusions. I will consider the solutions proposed in the literature to the problem of giving a semantics to underspecified representations, and conclude that these proposals either do not capture our intuitions about ambiguity, or result in inference patterns that can only be interpreted syntactically. I will then reconsider what we know about disambiguation, and suggest a way of looking at underspecification that doesn't suffer from either problem. Furthermore, the form of underspecification I will propose results in a much more general theory of ambiguity processing, one that also covers syntactic ambiguity.

2 Disambiguation as defeasible reasoning with underspecified representations

In this section I will briefly describe how underspecified representations are used to formulate theories of disambiguation using defeasible reasoning. I will introduce for this purpose a much simplified form of underspecified representation, that captures the essential spirit of the main existing proposals, but has just the minimum set of features necessary to make the problem I intend to address understood.

I call the underspecified language I am going to use SIMPLEST UNDERSPECIFIED LANGUAGE, or $\mathcal{SUL}$. From a syntactic point of view, existing underspecified language can be characterized as including two forms of expressions: 'disambiguated' or 'definite' expressions, corresponding to the expressions usually used in semantics to represent an interpretation; and 'underspecified' expressions, used to represent those aspects of interpretation that have to be resolved. $\mathcal{SUL}$ includes:

- **Definite Terms:** these are the constants k, h

[1] According to some authors, underspecified representations have a role in human cognitive processing that goes beyond disambiguation. I will not be concerned here with the question of when and whether humans disambiguate.

- **Underspecified Terms:** $\dot{x}$, $\dot{y}$, $\cdots$
- **Noun Phrases:** $[_{\mathrm{NP}}\ \alpha]$, for α a definite or underspecified term.
- **Definite Predicates:** the predicates croak_f, croak_d, and left.
- **Underspecified Predicates:** croak_U
- **Verb Phrases:** $[_{\mathrm{VP}}\ \alpha]$, for α a definite or underspecified predicate.
- **Basic Formulas:** all expressions of the form P(t), for t a term and P a predicate.
- **Structurally Underspecified Formulas:** all expressions of the form $[_{\mathrm{S}}\ \alpha\ \beta]$, for α a noun phrase and β a verb phrase.
- **Formulas:** Φ, for Φ a basic formula or an underspecified formula. If Φ and Ψ are formulas, then $\neg\Phi$ and $\Phi \wedge \Psi$ are formulas.

A language like the one just described can be used to express underspecified interpretations of sentences that are either lexically ambiguous (i.e., they contain an ambiguous lexical item, such as *bank*) or referentially ambiguous (i.e., they contain anaphoric expressions). The former are represented by underspecified predicates like croak_U, the latter by underspecified terms like $\dot{x}$. For example, the sentence *Kermit croaked*, containing the homonym *croak* which is ambiguous between a sense of 'uttering a frog-like sound' and 'dying', can be represented by the following structurally underspecified formula:

(1) $[_{\mathrm{S}}\ [_{\mathrm{NP}}\ k]\ [_{\mathrm{VP}}\ \mathsf{croak}_U]]$

whereas the sentence *he left* can be assigned, prior to resolving the anaphoric expression *he*, the following underspecified formula as its logical form:

(2) $[_{\mathrm{S}}\ [_{\mathrm{NP}}\ \dot{x}]\ [_{\mathrm{VP}}\ \mathsf{left}]]$

The language can also be used to represent the disambiguated interpretations of these sentences. The two predicates croak_f and croak_d represent the disambiguated readings of *croak*, so that, for example, the disambiguated interpretation of *Kermit croaked* according to which Kermit is dying, can be represented by the $\mathcal{SUL}$ expression $\mathsf{croak}_d(k)$. Similarly, the disambiguated interpretation of *he left* according to which it is Hobbes who left can be represented by the expression $\mathsf{left}(h)$.

Perhaps the best known feature of current underspecified languages is their capability of capturing the meaning of scopally underspecified interpretations. Noun Phrases, Verb Phrases and Structurally Underspecified Formulas have been included in the definition of $\mathcal{SUL}$ to show that this could be done in this language as well, once it were augmented with expressions that could be used to represent quantification. For example, if we added to the language the expressions every and a, we could represent the sentence *every kid climbed a tree* as:

(3) $[_{\mathrm{S}}\ [_{\mathrm{NP}}\ \mathsf{every(kid)}]\ [_{\mathrm{VP}}\ [_{\mathrm{V}}\ \mathsf{climbed}]\ [_{\mathrm{NP}}\ \ \mathsf{a(tree)}]]]$

But although scopal ambiguity has been the primary motivation for introducing underspecified representations, adding quantifiers would only make the discussion of the semantics of the language in the rest of the paper more complicated, without adding anything to the point I intend to discuss; therefore I will limit myself in this paper to a discussion of lexical and referential ambiguity.

The one known limit of languages like $\mathcal{SUL}$ is to express *syntactic* underspecification, e.g., the fact that *John saw the man with a telescope* can have two interpretations. This is not typically considered a problem, on the assumption that syntactic ambiguity is treated differently; the result however is that underspecified representations like $\mathcal{SUL}$ do not integrate well with models of disambiguation in which semantic plausibility checking is performed in parallel with syntactic structure building. We will return to this point again.

I will assume here that disambiguation is a process of defasible inference, and I will use Default Logic (Reiter, 1980) to formalize the reasoning involved in disambiguation. Default Logic is much simpler than the theories of nonmonotonic reasoning actually used in the NLP literature–among other things, there is no way of expressing priorities among defaults—but its properties are better known.

A quick reminder of Default Logic terminology: a DEFAULT INFERENCE RULE is a generalization of a traditional inference rule. It takes the form $\alpha(\mathrm{x}):\beta_1(\mathrm{x}) \ldots \beta_n(\mathrm{x})/\ \gamma(\mathrm{x})$, where $\alpha(\mathrm{x})$ is the precondition, $\gamma(\mathrm{x})$ is the consequent and $\beta_1(\mathrm{x}) \ldots \beta_n(\mathrm{x})$ is the set of JUSTIFICATIONS. For example, the default *Birds typically fly* is represented by the default rule **BIRD-FLY**,

(4) $$\frac{\mathsf{bird}(x) : \mathsf{flies}(x)}{\mathsf{flies}(x)}\ \textbf{BIRD-FLY}$$

whose interpretation is as follows: if x is a bird, and it is consistent to assume that x flies, then we can assume that x flies. A DEFAULT THEORY is a pair $\langle D, W\rangle$ where W is a set of formulae of the language L, and D a set of default inference rules. Whereas in normal logic such a set of formulas together with a consistent set of axioms and inference rules generates a single consistent theory, a default theory may generate more than one EXTENSION; each extension represents a consistent 'closure' of the set of formulae W under the logic of L augmented by a subset of the default inference rules in D.

Disambiguation can be modeled using a language like $\mathcal{SUL}$ and a nonmonotonic logic such Default Logic as an inference process which starts from a set of formulas possibly containing underspecified expressions and ends up with one (or more) extensions, each consisting of completely disambiguated expressions. Disambiguation rules can be formulated as default inference rules whose premisses are underspecified or definite formulas, and whose consequences are more specified formulas. These rules are mixed with default rules like **BIRDS-FLY** that encode commonsense knowledge. As an example, the process of lexical disambiguation may consist of rules such as:

(5) $$\frac{[\mathrm{S}\ [\mathrm{NP}\ \mathrm{x}]\ [\mathrm{VP}\ \mathsf{croak}_U]] \wedge \mathsf{frog}(x) :\ \mathsf{croak}_f(x)}{\mathsf{croak}_f(x)}\ \textbf{CROAK}_f\textbf{-IF-FROG}$$

CROAK$_f$-IF-FROG states that if the sentence [S [NP x] [VP croak_U]] is in the initial interpretation, and if it known that x is a frog, then assume that the intended interpretation of croak_U was that x uttered a frog-like sound, if it is consistent to do so.[2]

The process of disambiguating the sentence *Kermit croaked* can now be formalized as the process of computing the extensions of the default theory

(6) $$(\{\textbf{CROAK}_f\textbf{-IF-FROG}\},\ \{[\mathrm{S}\ [\mathrm{NP}\ \mathrm{k}][\mathrm{VP}\ \mathsf{croak}_U]],\ \mathsf{frog}(k)\})$$

that should have the single extension in (7), if you set up your logic right:

(7) $$\{[\mathrm{S}\ [\mathrm{NP}\ \mathrm{k}][\mathrm{VP}\ \mathsf{croak}_U]],\ \mathsf{frog}(k),\ \mathsf{croak}_f(k)\}$$

The machinery introduced so far can be used in a similar way to formalize the process of reference resolution and, given that the language is properly extended as discussed above, to formalize scope disambiguation. (For an example of how to do the former, see (Alshawi, 1992); of the latter, see (Poesio, 1994).)

This way of formalizing disambiguation has many advantages, most important of which is that it seems consistent with what is known about the way humans process ambiguity. Many recent psychological studies[3] seem to indicate that disambiguation involves generating multiple hypotheses in parallel, choosing the one which is most highly ranked if there is one such, and otherwise arriving at a PERCEIVED AMBIGUITY. This is what one would expect if disambiguation were a defeasible reasoning process with conflicting defaults; each alternative hypothesis would then correspond to a separate extension of the default theory. For example, if we were to add to the theory in (6) the new default inference rule in (8):

(8) $$\frac{[\mathrm{S}\ [\mathrm{NP}\ \mathrm{x}]\ [\mathrm{VP}\ \mathsf{croak}_U]] \wedge \mathsf{human\text{-}like}(x) :\ \mathsf{croak}_d(x)}{\mathsf{croak}_d(x)}\ \textbf{CROAK}_d\textbf{-IF-FROG}$$

we would obtain two extensions (see (Poesio, 1995) for discussion).

1. {[S [NP k][VP croak_U]], frog(k), human-like(k), $\mathsf{croak}_f(k)$}
2. {[S [NP k][VP croak_U]], frog(k), human-like(k), $\mathsf{croak}_d(k)$}

3 Reasoning with Underspecified Representations?

Both monotonic and nonmonotonic reasoning are built on the same assumptions about inference: we conclude that P is true if it can be derived from other wffs already known to be true. Thus, if we want to model disambiguation as a defeasible inference process, we have to say that some initial facts about the sentence are true, and to defeasibly infer a conclusion from these. As it turns out, the existing theories about the semantics of underspecified representations can be divided in two camps: those that preserve this traditional notion of

[2] The 'theory' of lexical disambiguation encoded by **CROAK$_f$-IF-FROG** should not be taken too seriously. This example is just meant to illustrate how disambiguation with underspecified representations is typically formulated.

[3] See, e.g., (Marslen-Wilson, 1975) for lexical disambiguation.

'inference' but do not really capture our intuitions about what ambiguity is, and those that result in a very different notion of 'inference'.

Underspecified representations are used to interpret semantically ambiguous expressions,[4] and more or less all attempts at providing a definition of semantic ambiguity try to capture the intuition that a semantically ambiguous expression is one that has 'multiple meanings'. One of the best attempts at providing a treatment of ambiguity is (Pinkal, 1995a). In that book, Pinkal makes ambiguity a special case of INDEFINITENESS: A sentence is SEMANTICALLY INDEFINITE if and only if in certain situations, despite sufficient knowledge of the relevant facts, neither "true" nor "false" can be clearly assigned as its truth value (Pinkal, 1995a). The notion of 'indefiniteness' is formalized in terms of the notion of PRECISIFICATION:

Precisification Principle: A sentence is of indefinite truth value in a situation if and only if it can be precisified alternatively to "true" or to "false".

where expression α in context c CAN BE PRECISIFIED to s iff (i) s is a sense that α can assume according to its meaning; and (ii) s is more precise than the sense of α in c.

Essentially all current theories of underspecification can be reconduced to this idea of precisification; where they differ is in how the notion of precisification is formalize. One approach to underspecification is to take the precisifications to be the disambiguated interpretations that can be extracted from the initial underspecified representation by means of syntactic operations on the initial underspecified logical form, such as replacing all underspecified predicates by disambiguated ones or, in the case of scopally ambiguous expressions, 'pulling out' operators by methods similar to those used by Hobbs and Shieber (1987). The original QLF proposal and UDRS proposal fall in this category (Alshawi and Crouch, 1992; Reyle, 1993).[5]

Pinkal himself and Deemter (van Deemter, 1991) hold a different view on precisifications. They both argue that what gets 'precisified' is the way that an expression gets interpreted, i.e., the function that is used to assign a value to logical expressions. Pinkal uses a variant of the supervaluation method, in which the value assigned by an interpretation function $[\![.]\!]$ to an expression α, $[\![\alpha]\!]$, is determined by the values assigned to that expression by a set $\mathcal{V}$ of 'classical' interpretation functions: in the case of a propositional expression p, for example, $[\![p]\!] = 1$ if for all V in $\mathcal{V}$, V(p) = 1; $[\![p]\!] = 0$ if for all V in $\mathcal{V}$, V(p) = 0; otherwise, $[\![p]\!]$ is undefined. The model used by Pinkal imposes the further assumption that a partial order relation of 'more precise than' is defined over

[4]The term 'underspecified' is itself ambiguous, and has also been used to indicate 'lacking information'. I will only use it to capture the form of ambiguity that in the case of lexical ambiguity goes under the name of homonimy.

[5]A variant of this approach, proposed in (Poesio, 1995), is to derive the set of precisifications 'semantically' or 'bottom-up', by making the expressions of the language denote sets of intensional objects (e.g., sets of functions from situations to truth values in the case of sentential expressions), and by letting the ambiguity of expressions 'propagate' so that, for example, an expression like $[_S\ [_{NP}\ k][_{VP}\ \mathsf{croak}_U]]$, which contains a sub-expression croak_U which denotes a set of predicate denotations, will in turn denote a set of propositions.

the interpretation functions in $\mathcal{V}$. This technique for assigning a semantics to underspecified expressions has very interesting logical properties, but is more limited than the other approach in that it does not extend easily to assign a semantics to scopally underspecified expressions.

Let us now abstract away from the technical implementation of the notion of precisification, and consider how we can define a notion of truth in an underspecified language, TRUE_U, given a treatment of semantic ambiguity in terms of precisification and a notion of truth in a standard logic 'TRUE'.[6] Let A be an underspecified expression, and let $\{A_1 \ldots A_n\}$ be its precisifications. Pinkal's Precisification Principle leads naturally to what Deemter called the 'strong' notion of truth: A is TRUE_U (FALSE_U) iff all of the precisifications A_j are TRUE (FALSE). The only alternative to be seriously examined in the literature is what has been called the 'weak' notion of truth: A is TRUE_U iff at least one of the precisifications A_j is TRUE.

The problem with the weak notion of truth is that it doesn't capture our intuitions about ambiguity. Consider the sentence *He left*, and consider a model in which there are only two objects, Kermit and Hobbes. According to the weak notion of truth, the sentence is true iff either Kermit or Hobbes left. But this is not what our intuitions seem to suggest: we want to say that the sentence is true iff the intended interpretation of it is true, and false otherwise. If speaker A utters the sentence intending to refer to Kermit, and Kermit stayed, we don't want to predict that the sentence was true in case Hobbes happened to leave.

The weak notion of truth does, however, have an advantage, and that is that there are no truth gaps. If we adopt the strong notion of truth, our disambiguation inferences take a very strange character. Consider the inference steps taken in the example of lexical disambiguation discussed in the previous section. They went as follows:

1. $[_S\ [_{NP}\ k][_{VP}\ \mathsf{croak}_U]]$ (premise)
2. $\mathsf{frog}(k)$ (premise)
3. $\mathsf{croak}_f(k)\}$ (1., 2., **CROAK$_f$-IF-FROG**)

If we assume the strong notion of truth the first premise, 1., is neither true nor false. So what we have is a sequence of inference steps which do not depend on the truth or falsity of the intermediate steps, but only on their syntactic form.

One could reply that there is nothing wrong with this: inference is typically defined in a syntactic fashion, and the notion of extension in Default Logic is even more syntactic that corresponding notions in other non-monotonic theories. Instead, what I am going to suggest is that we reconsider what it is that we are reasoning about when disambiguating; and that by doing so, we can get a theory of ambiguity and underspecification that does not rely on such inferences, and one, moreover, that is much more general than those proposed until now.

[6]Because I am mostly interested in defeasible inference, I will not consider the definition of entailment. For various ways of defining that, see again Deemter's thesis.

4 Disambiguation and the Discourse Situation

There is a widespread consensus that to disambiguate humans take into account both information about the semantic content of utterances and information about the DISCOURSE SITUATION—pragmatic facts about the context of utterance, each other's intentions, and so forth. An example of pragmatic information that clearly plays a role in disambiguation is information about the 'rhetorical structure' of a text or dialogue, defined as the way the intentions / speech acts expressed by each utterance relate to each other. This structure plays a role in the interpretation of definite descriptions and other anaphoric expressions (Grosz, 1977; Fox, 1987; Grosz and Sidner, 1986), of tense (Lascarides and Asher, 1991), and of ellipsis (Kehler, 1993). If we want to preserve Clark and Marshall's (1981) original hypothesis that that utterances are felicitous insofar as the speaker only relies on shared knowledge for their interpretation, we must conclude that the common ground consists of, or at least includes, information about the discourse situation, such as what speech acts were performed by each conversational participant, and their relations.

I propose to take the claim that the common ground includes information about the speech acts that took place as the foundation for my theory of underspecification. I want to go one step further than existing theories of speech acts, however, and claim that the common ground includes not only information about 'complete' speech acts, but also information about the occurrence of smaller utterance 'events,' such as the event of uttering a single word, or even a fragment of a word. I will call these 'small utterance events' MICRO CONVERSATIONAL EVENTS. I propose, that is, that as a result of an utterance of the sentence *Kermit croaked*, we obtain a common ground that includes the information in (9).

(9)

μ-ce$_1$,μ-ce$_2$,μ-ce$_3$
...
μ-ce$_1$: $\boxed{\mathsf{utter}(a, \textit{“kermit”})}$
μ-ce$_1 \rightsquigarrow k$
$\mathsf{cat}(\mu\text{-ce}_1) = \text{NP}$
μ-ce$_2$: $\boxed{\mathsf{utter}(a, \textit{“croaked”})}$
$\mathsf{cat}(\mu\text{-ce}_2) = \text{VP}$
$\mathsf{cat}(\mu\text{-ce}_3) = \text{S}$
$\mu\text{-ce}_2 \prec \mu\text{-ce}_3$
$\mu\text{-ce}_1 \prec \mu\text{-ce}_3$
...

We have in (9) three 'micro-conversational events': an event μ-ce$_1$ of uttering the word *Kermit*, of category NP; an event μ-ce$_2$ of uttering the word *croaked*, of category VP; and an event μ-ce$_3$ of syntactic category S, with constituents μ-ce$_1$ and μ-ce$_2$.

I have used a DRT-like 'box' notation, but I will use this representation as a quasi-first order language with the usual semantics, in which terms denote

objects in the universe, and properties denote sets of objects. The 'discourse referents' μ-ce$_1$, μ-ce$_2$, and μ-ce$_3$ can be thought of as constants. The language includes the following function symbols:

- the function cat denotes the function that maps a micro-ce into its syntactic category.
- the symbol $\rightsquigarrow$ denotes the (family of) functions that map a micro-conversational event into their CONTENT (see below); I write this function in infixed form.
- the symbol $\prec$ denotes the (partial) function that maps a micro-conversational event into the micro-conversational event it is a constituent of. I will also use the symbol $\prec^*$ to indicate the transitive closure of $\prec$, i.e., $e \prec^* e'$ iff either $e = e'$ or there is an e'' such that $e \prec^* e''$ and $e'' \prec^* e'$.

Events will be treated in the usual Davidsonian manner, according to which predicates have an extra argument for the event. I represent the Davidsonian argument 'outside' the predicate proper, as done in (Kamp and Reyle, 1993) — for example, the event of Mary killing John will be represented as in e:$\boxed{\mathsf{kill}(m,j)}$.

The key properties of micro-conversational events are (i) the existence of a function $\rightsquigarrow$ from utterance events to their CONTENT, an object denoting their ordinary semantic translation; and (ii) that the content may be at some point unspecified. What makes the common ground represented in (9) 'underspecified' is that the contents of μ-ce$_2$ and μ-ce$_3$ are not provided. This common ground only specifies about μ-ce$_2$ that an event of uttering the phonetic string "croaked" took place; what content this event had is left unspecified. This is, I suggest, the underspecified representation of lexical items, prior to lexical access. The kind of underspecification involved in lexical ambiguity is interpreted here as lack of information about the content of a micro conversational event of uttering a lexical item; and the (lexical) disambiguation task is formulated as the task of inferring the content of micro-conversational events. As we will see below, this model can also be used to capture a second form of underspecification, that we may call 'structural'.

Let us see how the underspecified representation I propose can serve as the basis for formalizations of disambiguation analogous to those seen in section §2. Let me introduce the following abbreviations:

- μ-ce$_1$:[NP :sem k] stands for $\mathsf{cat}(\mu\text{-ce}_1) = \text{NP}$ and $\mu\text{-ce}_1 \rightsquigarrow k$
- μ-ce$_2$:[VP :phon "croaked"] stands for:
  $\mathsf{cat}(\mu\text{-ce}_2) = \text{VP}$
  μ-ce$_2$: $\boxed{\mathsf{utter}(a, \textit{“croaked”})}$
- μ-ce$_3$: [S μ-ce$_1$:α μ-ce$_2$:β] stands for:
  $\{\ \mathsf{cat}(\mu\text{-ce}_3) = \text{S},$
  $\mu\text{-ce}_2 \prec \mu\text{-ce}_3,$
  $\mu\text{-ce}_1 \prec \mu\text{-ce}_3\} \cup$
  $\{\ \mu\text{-ce}_1{:}\alpha,\ \mu\text{-ce}_2{:}\beta\}$

the rule **CROAK$_f$-IF-FROG** can now be reformulated as follows:

$$\frac{\mu\text{-}ce_3{:}[_S\ \mu\text{-}ce_1{:}\ [_{\text{NP}}\ \text{:sem}\ x]\ \mu\text{-}ce_2{:}\ [_{\text{VP}}\ \text{:phon}\ \text{“croaked”}]] \wedge \mathsf{frog}(x) \quad : \mu\text{-}ce_2 \rightsquigarrow \mathsf{croak}_f}{\mu\text{-}ce_2 \rightsquigarrow \mathsf{croak}_f}\ \textbf{CROAK}_f\textbf{-IF-FROG}$$

the default theory consisting of (9) and of this default rule is:

(10) $(\{\textbf{CROAK}_f\textbf{-IF-FROG}\},$
$\{\mu\text{-}ce_3{:}[_S\ \mu\text{-}ce_1{:}\ [_{\text{NP}}\ \text{:sem}\ k]\ \mu\text{-}ce_2{:}\ [_{\text{VP}}\ \text{:phon}\ \text{“croaked”}]], \mathsf{frog}(k)\})$

This default theory has the extension in (11).

(11) $\{\mu\text{-}ce_3{:}\ [_{\text{S}}\ \mu\text{-}ce_1{:}[_{\text{NP}}\ k]\ \mu\text{-}ce_2{:}[_{\text{VP}}\ \mathsf{croak}_f]\ ], \mathsf{frog}(k)\}$

The crucial point, with respect to the discussion in the previous sections, is that (11) is obtained through a fairly standard (if defeasible) chain of inference steps. The underspecified language just proposed, unlike $\mathcal{SUL}$, does not require a special semantics, since it does not contain constructs equivalent to those used in $\mathcal{SUL}$ to represent underspecification. The functions $\prec$ and $\rightsquigarrow$ can be given an interpretation within an ordinary (typed) model. This is also true of the construct used to represent events. Hence, (9) is going to represent either a true or false state of affairs. Our inference chain starts from what we know is true (about the discourse situation), and we hypothesize that something else is true.

What is, then, the connection between this form of underspecification and Pinkal's theory of 'ambiguity as multiplicity of precisifications'? Where did the notion of 'precisification' go? The answer is that the model-theoretic notions of precisification discussed in previous sections are going to be replaced by a proof-theoretic one. The precisifications of *Kermit croaked* in a determined discourse situation, I suggest, are the extensions generated by the default theory that formalizes that discourse situation. This move has at least two significant consequences. First of all, we give up on an 'absolute' notion of ambiguity: our notion of 'ambiguity' is relativized to discourse situations, in the sense that the same expression is going to have different precisifications in different contexts. Secondly, controlling the combinatorial explosion of interpretations becomes a matter of controlling the inferences.

Some readers may perhaps wonder whether it was worth our while to come up with yet another theory about underspecification, if the only advantage of this theory were that it preserves the traditional notion of inference. In fact, the most important advantage of this theory is not that allows us to formalize disambiguation within a traditional semantics, but that it leads to a more general and, in many ways, more plausible account of ambiguity and of the disambiguation process. We will see in the next section a clear case of disambiguation that can be formalized within this framework, but not using languages like $\mathcal{SUL}$. But the advantages of the current proposal over $\mathcal{SUL}$-style approaches to underspecification can already be seen with lexical disambiguation. The theory of lexical underspecification and disambiguation discussed in the previous sections

depended on the assumption that the lexical entry of *croak* referred to a predicate croak_U. This is a pretty implausible hypothesis: under most theories of lexical ambiguity, homonym words are distinguished from polysemous words in that the former, unlike the latter, have distinct lexical entries (see, e.g, (Cruse, 1995). In the framework just proposed, lexical disambiguation can be formalized in a much more plausible way, by treating each lexical entry for the word-string *croak* as a default triggered by the occurrence of a micro-conversational event of uttering that word string; 'underspecified' predicates such as croak_U can be reserved for polysemous words such as *mouth* or *newspaper*, in which case they can be given a standard predicate semantics. In other words, the framework just introduced is much more consistent with what is currently known about lexical ambiguity.

5 Syntactic Ambiguity, Robust Discourse Interpretation, and Radical Underspecification

Theories of underspecification based on languages like $\mathcal{SUL}$ cannot be used to formalize syntactic ambiguity. The model just discussed, on the contrary, generalizes to the case of syntactic ambiguity in a straightforward fashion. A 'radically underspecified' interpretation of *Kermit croaked*, i.e., one in which not even the syntactic structure of the sentence has been determined,[7] is simply one in which the information about sub-constituents is not present:

(12)

$\mu\text{-}ce_1, \mu\text{-}ce_2, \mu\text{-}ce_3$
...
$\mu\text{-}ce_1$: $\boxed{\mathsf{utter}(a, \text{“}kermit\text{”})}$
$\mu\text{-}ce_1 \rightsquigarrow k$
$\mathsf{cat}(\mu\text{-}ce_1) = \text{NP}$
$\mu\text{-}ce_2$: $\boxed{\mathsf{utter}(a, \text{“}croaked\text{”})}$
$\mathsf{cat}(\mu\text{-}ce_2) = \text{V}$
...

The common ground in (12) contains just the information that a noun phrase and a verb were uttered. If we take (12) as our initial interpretation of *Kermit croaked*, any other information about syntactic structure would have to be inferred in the course of disambiguation, by means of default rules such as (1), that infers the occurrence of an event of category 'S' when an event of type 'NP' is immediately followed by an event of cat 'VP'. (μ-ce$_3$ is a new micro conversational event.)

The view of disambiguation defended in this paper has obvious connections with the 'parsing as deduction' perspective on syntactic disambiguation taken in Categorial Grammar. In particular, the inferences deriving new sentential

[7] The term 'radical underspecification' is due, as far as I know, to Manfred Pinkal (p.c.), who is also developing an approach to underspecification capable of dealing with the issues discussed in this section (Pinkal, 1995b).

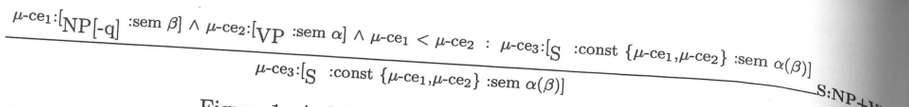


Figure 1: A defeasible rule for parsing

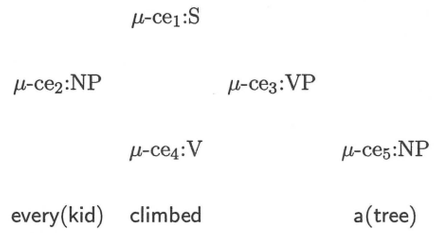
constituents in Categorial Grammar can be reanalyzed within the current model of disambiguation as inferences hypothesizing the occurrence of larger micro-conversational events—with the difference, of course, that the nice computational properties of categorial grammar are lost if we reinterpret the constructors as defaults. More in general, a parser that works by defeasible reasoning over underspecified representations, rather than by the tried and tested (and more efficient) methods found in the literature, may be taken as an overkill. will not be taken very seriously. I am not suggesting that we should throw away the existing techniques for treating parsing ambiguities in an efficient fashion, and replace them with defeasible reasoning processes of high complexity. Yet, there are some reasons for thinking of syntactic interpretation rules as defaults that operate together with other defaults. One of them is that this may explain the interactions between pp-attachment and reference resolution observed, for example, by Crain and Steedman (1985). A second reason is found by observing the kind of input one sees in spoken conversations. Consider for example the fragment in (13), from the 1991 TRAINS transcripts collected at the University of Rochester:

(13)
```
88.3    there's
88.4    let me
88.5    let me summarize this and
88.6    make sure I did everything correctly
88.7    okay
88.8    the current plan is
88.9    we take
88.10   f okay
88.11   I'll say
88.12   we'll start with the bananas
88.13   we take engine E2
88.14   ... and
```

Current theories about how to deal with this kind of input usually involve one of two strategies: either to claim that parsing doesn't really involve any syntactic knowledge at all; or to assume a 'dual parser' strategy, in which a syntactically motivated parser and a semantically-driven parser work in tandem, and some criterion exist for choosing between the outputs of the two. The model of radical underspecification just proposed suggests a third possibility: that spoken input is first processed by a lexical processor and maybe a partial parser such as Fidditch (Hindle, 1990), resulting in an radically underspecified interpretation as in (12); and that the subsequent steps of disambiguation are performed by default rules with different priorities, which encode ways of combining micro-conversational events both according to syntactic knowledge and according to semantic knowledge.

6 Scope Ambiguity

So far I haven't said anything about scope ambiguity, which is the kind of ambiguity that historically has motivated the work on underspecified representations. It should be clear that once we can give an interpretation to structurally ambiguous sentences, we can also give an interpretation for scopally ambiguous ones: the discourse situation in the case of a scopally ambiguous sentence such as *Every kid climbed a tree* will be as in (14). (I have used the abbreviations introduced in previous sections.)

(14)


The really challenging task, then is to show how scope disambiguation can work within this framework. Most models of scope disambiguation assume the existence of an inference mechanism able to infer scope relations between operators: thus, in Reyle's Underspecified DRS, in which every DRS 'box' is assigned a label, it is assumed that disambiguation is performed by unconvering ordering relations $l_i \leq l_j$ among labels; in (Poesio, 1991), as in QLF (Alshawi and Crouch, 1992), an index is assigned to each operator, and underspecified representations are augmented with sets of ordering constraints among these indices. In all of these proposals, the ordering among labels affects the semantics of the underspecified expression, in the sense that the procedures that compute the set of precisifications of underspecified expressions only generate the readings that do not violate the orderings. Partial as well as total disambiguations can be represented. I will show that this theory of scopal disambiguation can be modeled faily easily within the theory of underspecification being proposed.

There is an obvious connection betweeen this way of thinking about partial scope disambiguation and Gawron and Peters' proposal (1990) about the way scope information is encoded in the discourse situation (that they call 'circumstances of interpretation'), and about the effect of this scope information on semantic interpretation. In Gawron and Peters' theory of grammar, the information contained in the discourse situation affects the rules of semantic composition. The part of their proposal that is relevant to our discussion is their assumption that the discourse situation encodes information about the relative scope of utterances of noun phrases, which they represent by means of relations like scopes-over(*"every kid"*, *"a tree"*); and that this information plays a role in determining the semantic interpretation of an utterance. According to Gawron and Peters, for example, the semantic interpretation of the sentence *Every kid climbed a tree* will be one in which the quantifier every(kid) takes wide scope over the quantifier a(kid) in a discourse situation which contains the information that scopes-over(*"every kid"*, *"a tree"*); whereas in a discourse sit-

uation which contains information that scopes-over(*"a tree"*, *"every kid"*), the same sentence will have the other disambiguated interpretation.

The scopes-over relation obviously encodes the same form of information that is encoded by relations among labels like $l_i \leq l_j$ in theories such as UDRS. Gawron and Peters' proposal thus suggests a way in which information about partial disambiguations may be encoded in the micro-conversational events model—i.e., by means of relations of the form scopes-over(μ-ce$_i$,μ-ce$_j$). Their ideas about the role of scope information on semantic interpretation also suggest a way of looking at how this information about partial disambiguation affects the choice of a precisification that is in line with the view of precisification adopted in this paper, according to which the precisifications of an utterance correspond to the extensions of a default theory encoding the information in the discourse situation that contains that utterance.

The idea I want to pursue is that facts like scopes-over(μ-ce$_i$,μ-ce$_j$) play a role in the default rules that specify how micro conversational events combine into larger units. This idea can be formulated most perspicuously by adopting the view of the connection between grammar and scope assignment developed in recent proposals in the Categorial Grammar framework, in particular (Pereira, 1990; Moortgat, 1990; Carpenter, 1994; Carpenter, 1996). In this work, the tight connection between operator scopes and derivations adopted in Montague Grammar is resurrected. This version of categorial grammar includes, in addition to the usual inference schemes, a SCOPING CONSTRUCTOR that accounts for the way quantified NPs participate to the construction of the meaning of a sentence. These authors implement the idea that the translations of generalised quantifiers and proper names are related by raising and lowering operations (Partee and Rooth, 1983; Hendriks, 1987) by introducing two distinct syntactic categories for noun phrases: one for proper names, NP, and a 'raised' one S $\Uparrow$ NP for generalised quantifiers. The scoping constructor allows 'lowering' the raised translation S $\Uparrow$ NP to one of type NP that can participate in a normal derivation, and subsequently 'raising' the translation again.[8] The natural deduction formulation of the scoping scheme is as follows:

Scoping Natural Deduction Scheme :

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \alpha : S \Uparrow NP\end{array}}{x : NP}\Uparrow E^n \quad \begin{array}{c}\vdots\\ \beta : NP\end{array}}{\alpha(\lambda x.\beta) : S}\, n$$

In Figure 2 is an illustration of how the Scoping Scheme is used in a derivation (of *every kid climbed a tree*) and how the scheme interacts with other construc-

[8] As Carpenter observes, the first part of the scheme corresponds to a storage operation in Cooper's account of scope (Cooper, 1983), while the second corresponds to a discharge operation.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\textit{every kid}}D}{\mathsf{every(kid)} : S \Uparrow NP}\Uparrow E^1}{x_1 : NP} \quad \dfrac{}{\mathsf{climbed} : NP\backslash S/NP}Lx \quad \dfrac{\dfrac{}{\textit{a tree}}D}{\mathsf{a(tree)} : S \Uparrow NP}\Uparrow E^2}{\mathsf{climbed}(x_1, x_2)}D}{x_2 : NP}D}{\mathsf{climbed}(x_1, x_2)}2}{\mathsf{a(tree}, \lambda x_2.\ \mathsf{climbed}(x_1,x_2))}3}{\mathsf{every(kid}\ , \lambda x_1.\ \mathsf{a(tree}, \lambda x_2.\mathsf{climbed}(x_1,x_2)))}$$

Figure 2: An example of use of the Scoping Scheme

tors. Both *every kid* and *a tree* are interpreted as generalised quantifiers, i.e., as expressions of type S $\Uparrow$ NP. The first part of the scheme—the 'lowering' or 'storage' operation—is applied to obtain out of S $\Uparrow$ NP's NPs that can participate in the derivation. At the end of the derivation, two 'discharge' operations are then performed; depending on the order of these operations, one or the other scope is derived. The derivation in Figure 2 shows how the interpretation with *every kid* taking wide scope is obtained; the other interpretation is obtained by reversing the order of application.

I suggested in the previous section that the inferences deriving new sentential constituents in Categorial Grammar can be reanalyzed within the current framework as default inferences hypothesizing the occurrence of larger micro-conversational events. Similarly, the Carpenter / Moortgat / Pereira view of how operators are assigned their scope can be adopted within our framework by reformulating their scoping scheme in terms of default inference rules; one of the advantages of doing so would be to inherit the interesting treatment of scope islands that comes with this proposal. (See (Carpenter, 1996).) This is is essentially what I am going to do, except that I propose to merge this view with Gawron and Peters' hypothesis about the effect of scopal information in the discourse situation on the derivation of a semantic interpretation, as follows.

The Scoping Scheme is reformulated in terms of two inference rules: one allowing to infer, from the utterance of a quantified NP, the utterance of a non-quantified NP, whose content is a variable; and one allowing to infer, from the occurrence of a micro-ce μ-ce$_3$ of cat S, derived by 'storing away' the quantified NP μ-ce$_1$, the occurrence of a new micro-conversational event μ-ce$_4$, of cat S, whose content is obtained by discharging the content of the quantified NP μ-ce$_1$. I use categories NP[-q] and NP[+q] rather than NP and S $\Uparrow$ NP for simplicity.

The crucial difference between the rules proposed here and the Carpenter / Pereira account is that the **DISCHARGE**$_i$ rule(s) are only licensed when the appropriate scopes-over facts have been inferred. In the absence of such information, these rules do not apply. The reading of *Every kid climbed a tree* in which *every kid* takes wide scope would be inferred by the derivation in

$$\frac{\mu\text{-ce}_1{:}[\text{NP[+q]}\ \text{:sem}\ \beta]\ :\ \mu\text{-ce}_1{:}[\text{NP[-q]}\ \text{:sem}\ x_i]}{\mu\text{-ce}_1{:}[\text{NP[-q]}\ \text{:sem}\ x_i]}\ \textbf{STORE}$$

$$\frac{\begin{array}{l}\mu\text{-ce}_1{:}[\text{NP[+q]}\ \text{:sem}\ \beta] \wedge \mu\text{-ce}_1 \prec \mu\text{-ce}_2 \wedge \\ \mu\text{-ce}_2{:}[\text{NP[-q]}\ \text{:sem}\ x_i] \wedge \mu\text{-ce}_2 \prec^* \mu\text{-ce}_3 \wedge \\ \mu\text{-ce}_3{:}[\text{S}\ \text{:sem}\ \alpha] \wedge \\ \text{scopes-over}(\mu\text{-ce}_1, \mu\text{-ce}_3)\end{array} \quad : \ \mu\text{-ce}_4{:}[\text{S}\ \text{:sem}\ \beta(\lambda x_i.\alpha)]}{\mu\text{-ce}_4{:}[\text{S}\ \text{:sem}\ \beta(\lambda x_i.\alpha)]}\ \textbf{DISCHARGE}_i$$

Figure 3: Storage and discharge rules

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\textit{every kid}}\,D}{\mu\text{-ce}_1{:}[\text{NP[+q]}\ \text{:sem every(kid)}]}}{\mu\text{-ce}_4{:}[\text{NP[-q]}\ \text{:sem}\ x_1]}\,\textbf{store} \qquad \dfrac{\dfrac{}{\textit{climbed}}\,Lx}{\mu\text{-ce}_2{:}[\text{V}\ \text{:sem climbed}]} \qquad \dfrac{\dfrac{}{\textit{a tree}}\,D}{\mu\text{-ce}_3{:}[\text{NP[+q]}\ \text{:sem a(tree)}]}\,\textbf{store}}{\mu\text{-ce}_5{:}[\text{NP[-q]}\ \text{:sem}\ x_2] \qquad \mu\text{-ce}_6{:}[\text{VP}\ \text{:sem climbed}(x_2)]}\,\textbf{VP:V+NP}}{\mu\text{-ce}_7{:}[\text{S}\ \text{:sem climbed}(x_1,x_2)] \wedge \text{scopes-over}(\mu\text{-ce}_3, \mu\text{-ce}_7)}\,\textbf{S:NP+VP}}{\mu\text{-ce}_8{:}[\text{S}\ \text{:sem a(tree,}\lambda x_2.\ \text{climbed}(x_1,x_2))] \wedge \text{scopes-over}(\mu\text{-ce}_1, \mu\text{-ce}_8)}\,\textbf{DISCHARGE}_2}{\mu\text{-ce}_9{:}[\text{S}\ \text{:sem every(kid ,}\lambda x_1.\ \text{a(tree,}\lambda x_2.\text{climbed}(x_1,x_2)))]}\,\textbf{DISCHARGE}_3$$

Figure 4: An example of scope disambiguation

Figure 4.[9]

7 Related Work

That ambiguity involves 'going meta' was already proposed by Nerbonne (Nerbonne, 1992) and Reyle (Reyle, 1993).[10] The following quote is from Nerbonne:

> ...semantics distinguishes two levels, a level of meaning specification [or 'metalinguistic', MP] and a level of meaning (or denotation) ... ambiguity is disjunction (or underspecification) at the first level ...

The theory of underspecification just proposed can be seen as a way of 'going meta' which involves minimum change from our current theories of disambiguation. Instead of introducing a new level of representation and a new semantics, we exploit the fact that information about speech acts has to be present in the discourse model anyway, and just add information about 'smaller' kinds of speech acts.

There also seem to be connections between this proposal and the recent work by Dalrymple and colleagues on 'meaning assembly' (Dalrymple et al., 1995), that is also inspired by work on categorial grammar. (The symbol chosen for the 'content' function is meant to emphasize this connection.) There also are obvious differences. The form of representation presented here extends the one used by Dalrymple and colleagues in that it does not presuppose the existence of a parse tree which the meaning interpretation 'builds on': whatever syntactic structure is present, is made explicit. In this way, the representation can also be used when no syntactic structure is obtained.

The second difference between this proposal and the work by Dalrymple et al. is in the reasoning mechanism that is assumed. A big part of their proposal is the idea of using linear logic, whose main property is that 'resources matter.' Linear logic has computational advantages with respect to default logic, so it would be interesting to investigate the relation between the two in some detail.[11]

Finally, those familiar with the AI literature may recognize a connection between the micro-conversational events idea and the 'flat' representations proposed by Hobbs (Hobbs, 1983; Hobbs, 1986). I believe that the current proposal may clarify the connection between the representations used by Hobbs and more traditional theories theories of semantic interpretation.

[9] The proposal about scope and scope disambiguation just discussed shares with all current approaches to underspecification the assumption that it is possible to infer information about the relative scope of operators, and that the interpretation of a sentence is obtained by accumulating such information until a complete order is obtained. Unfortunately, nobody has a clear idea about how this information could be inferred. For an alternative view of scope disambiguation, see (Poesio, 1994).

[10] Crouch similarly proposed in (1995) that underspecified representations should be taken as 'descriptions'.

Acknowledgments

This work was borne out of the many discussions on underspecification I have had with Robin Cooper, David Milward, and with the members of the FRACAS project, especially Dick Crouch and Manfred Pinkal. Robin's work on Situation Theoretic Grammar has been particularly inspiring. I would also like to thank Kees van Deemter, Alex Lascarides and John Nerbonne for suggestions and constructive criticism.

References

Alshawi, H., editor. 1992. *The Core Language Engine.* The MIT Press.

Alshawi, H. and R. Crouch. 1992. Monotonic semantic interpretation. In *Proc. 30th. ACL*, pages 32–39, University of Delaware.

Brown, P., S. Della Pietra, V. Della Pietra, and R. Mercer. 1991. Word sense disambiguation using statistical methods. In *Proc. of the 29th ACL*, pages 264–270.

Carpenter, B. 1994. Quantification and scoping: a deductive account. In *Proc. of the 13th West Coast Conference on Formal Linguistics.*

Carpenter, B. 1996. *Lectures on Type-Logical Semantics.* MIT Press.

Clark, H. H. and C. R. Marshall. 1981. Definite reference and mutual knowledge. In *Elements of Discourse Understanding.* Cambridge University Press, New York.

[11] In recent unpublished work, Crouch and von Genabith explore the possibility of using an approach like Dalrymple et al.'s to provide a foundation for underspecification.

Cooper, R. 1983. *Quantification and Syntactic Theory.* Dordrecht, Holland: D. Reidel.

Crain, S. and M. Steedman. 1985. On not being led up the garden path: the use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen, and A. M. Zwicky, editors, *Natural Language Parsing: Psychological, Computational and Theoretical perspectives.* Cambridge University Press, New York, pages 320–358.

Crouch, R. 1995. Ellipsis and quantification: A substitutional approach. In *Proceedings 7th Conference of the European Chapter of the Association for Computational Linguistics*, pages 229–236, Dublin. Dublin City University.

Cruse, D. A. 1995. Polysemy and related phenomena from a cognitive linguistic viewpoint. In P. Saint-Dizier and E. Vegas, editors, *Computational Lexical Semantics.* Cambridge University Press, chapter 2, pages 33–49.

Dalrymple, M., J. Lamping, F. Pereira, and V. Saraswat. 1995. Linear logic for meaning assembly. Unpublished manuscript circulated at the 1995 ESSLI.

Fox, B. A. 1987. *Discourse Structure and Anaphora.* Cambridge, UK: Cambridge.

Gawron, J. M. and S. Peters. 1990. *Anaphora and Quantification in Situation Semantics.* Lecture Notes, volume 19. CSLI.

Grosz, B. J. 1977. *The Representation and Use of Focus in Dialogue Understanding.* Ph.D. thesis, Stanford University.

Grosz, B. J. and C. L. Sidner. 1986. Attention, intention, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

Hendriks, H. 1987. Type change in semantics: the scope of quantification and coordination. in E. Klein and J. van Benthem (eds.), Categories, Polymorphism and Unification, Centre for Cognitive Science, University of Edinburgh.

Hindle, D. and M. Rooth. 1993. Structural ambiguity and lexical relations. *Computational Linguistics*, 19(1):103–120. Special Issue on Using Large Corpora, I.

Hindle, Donald. 1990. A parser for text corpora. Technical memorandum, AT&T Bell Laboratories.

Hobbs, J. R. 1983. An improper treatment of quantification in ordinary English. In *Proc. ACL-83*, pages 57–63, Cambridge, MA, June.

Hobbs, J. R. 1986. Discourse and inference. Unpublished draft, July.

Hobbs, J. R. and S. M. Shieber. 1987. An algorithm for generating quantifier scopings. *Computational Linguistics*, 13(1-2):47–63, January-June.

Hobbs, J. R., M. Stickel, P. Martin, and D. Edwards. 1990. Interpretation as abduction. Technical Note 499, SRI International, Menlo Park, CA, December.

Kamp, H. and U. Reyle. 1993. *From Discourse to Logic.* Dordrecht: D. Reidel.

Kehler, A. 1993. The effect of establishing coherence in ellipsis and anaphora resolution. In *Proc. 31st ACL*, pages 62–69.

Lascarides, A. and N. Asher. 1991. Discourse relations and defeasible knowledge. In *Proc. ACL-91*, pages 55–63, University of California at Berkeley.

Marslen-Wilson, W.D. 1975. Sentence perception as an interactive parallel process. *Science*, 189:226–228.

Moortgat, M. 1990. The quantification calculus: Questions of axiomatization. In DYANA Deliverable R1.2.A, ESPRIT Basic Research Action BR3175.

Nerbonne, J. 1992. A feature-based syntax/semantics interface. Research Report RR-92-42, DFKI, Saarbrücken, August.

Partee, B. H. and M. Rooth. 1983. Generalized conjunction and type ambiguity. In R. Bauerle, C. Schwarze, and A. von Stechow, editors, *Meaning, Use and Interpretation of Language.* Walter de Gruyter, Berlin, West Germany.

Pereira, F. C. N. 1990. Categorial semantics and scoping. *Computational Linguistics*, 16(1):1–10, March.

Perrault, C. R. 1990. An application of default logic to speech act theory. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication.* The MIT Press, Cambridge, MA, chapter 9, pages 161–185.

Pinkal, M. 1995a. *Logic and Lexicon.* Dordrecht: D. Reidel.

Pinkal, M. 1995b. Radical underspecification. In P. Dekker, J. Groenendijk, , and M. Stokhof, editors, *Proceedings of the Tenth Amsterdam Colloquium.*

Poesio, M. 1991. Relational semantics and scope ambiguity. In J. Barwise, J. M. Gawron, G. Plotkin, and S. Tutiya, editors, *Situation Semantics and its Applications, vol.2.* CSLI, Stanford, CA, chapter 20, pages 469–497.

Poesio, M. 1994. *Discourse Interpretation and the Scope of Operators.* Ph.D. thesis, University of Rochester, Department of Computer Science, Rochester, NY.

Poesio, M. 1995. Semantic ambiguity and perceived ambiguity. In K. van Deemter and S. Peters, editors, *Semantic Ambiguity and Underspecification.* CSLI, Stanford, CA.

Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence*, 13(1–2):81–132, April.

Reyle, U. 1993. Dealing with ambiguities by underspecification: Construction, representation and deduction. *Journal of Semantics*, 10:123–179.

van Deemter, K. 1991. *On the Composition of Meaning.* Ph.D. thesis, University of Amsterdam.

Woods, W. A. *et al.* 1972. The lunar sciences natural language information system: Final report. Report 2378, BBN, Cambridge, MA.

Is aspectual classification of nouns possible?*

E. Rakhilina (VINITI, Moscow)

Summary *A semantic classification of (Russian) nouns is proposed which corresponds to the verbal classification used in (Slavic) aspectological studies. This classification specifies the stability of a corresponding object with respect to the time axis. One interesting result of it (discussed in detail) is that it predicts the semantic interpretation of attributive constructions with the Russian "inner-temporal" adjective STARYJ 'old': this interpretation directly depends on the semantic class the noun belongs to. Thereby, in a sense, one can speak about aspectual types of Russian nouns.*

0. Semantic boundaries between nouns and verbs (especially, in a cross-linguistic perspective) are known to be not very rigid (the problem is discussed, e.g., in Givón 1979: 320-322; cf. also Bach 1968 for a logical aspect). In general, nouns and verbs seem to share much more properties than it is commonly thought. The most evident semantic zone is, in this respect, the temporal one. Indeed, it is well-known that not only verbs, but also nouns can take temporal characteristics. The grammatical tense relates the situation to the moment of speech, and this could be applied to certain nouns as well, including **concrete** nouns, and among them, first of all, names of kinds of persons (like *brother*, *professor* or *seaman*). It was shown, in particular, in [Dahl 1970], where these nouns were analyzed in the context of modifiers like *former* or *future*; there are also several works dealing with similar problems on Slavic and especially Russian material (for further references see [Jakovleva 1994]).

The idea is that the constructions like a *former champion* are interpreted in the same way as the verbs in the past tense. That is, a former champion is a person who WAS a champion in the PAST. The temporal reference is made here to the property of 'being a champion' expressed by the noun. We will speak in this case of **external** temporal characteristics of a verb or a noun. So *former* or *future* will be the external temporal modifiers.

Meanwhile, whenever we deal with the verbs, we must take into account the so-called **internal** temporal characteristics which are inherent to the meaning of the lexeme. According to the tradition, these characteristics are called **aspectual**. By means of different aspectual grammemes, the situation is distributed in time: it could be, for example, "condensed", or "prolonged",

or "multiplied", and so on. On the other side, it is known that the verbs are very idiosyncratic with respect to different aspectual characteristics: not each situation is capable of being, for example, prolonged or condensed. These facts explain why lexical classification of verbs is so crucial for aspectological theory (the famous article of Vendler [1967] is to be mentioned first in this connection).

Now, as we know, the aspectual characteristics can be easily ascribed also to the **deverbal** nouns. In Russian, we have a pair of a perfective verb *spasti* 'to save, to have saved' and an imperfective verb *spasat'* 'to save, to save usually', and each of them can form a corresponding deverbal noun denoting agent: *spasitel'* from *spasti*, *spasatel'* from *spasat'*. Thus, the pair *spasitel'* - *spasatel'* displays the aspectual opposition of *spasti* - *spasat'*.

But the question arises whether concrete nouns (names of objects) can display the same type of internal temporal characteristics. In what follows I will try to answer this question.

1. In order to do this, I have studied the use of Russian concrete nouns in the constructions with the adjective *staryj* 'old'. The analysis will begin with, so to speak, "old things" - i.e. constructions where a noun denotes a concrete object (cf. *old house*, *old tree*, and so on). In this area we can distinguish four types of interpretation which, as it will be shown later, correspond to the four distinct semantic classes of nouns [1].

These four classes can be exemplified by the following typical noun phrases:

1. *old oak*
2. old rag
3. *old channel*
4. old coins

Let us consider these classes in turn.

The semantic interpretation imposed on the noun by the adjective within the first class is 'the object that arose a long time ago and has existed for a long time' (further examples are: old *grove*, *elephant*, *mushroom*, *callosity*, etc.). Note that in some cases *young* could serve as a possible antonym (applicable, first of all, to animate nouns of this class).

The semantic interpretation within the second class is 'the object that, being in use for a long time, has become useless and/or decayed' (further

examples are: *old clothes*, *ladder*, *house*, *furniture*, etc.). A good antonym in most cases is *new*. The important peculiarity of this class is that the above interpretation usually coexists with others, and the overwhelming majority of the examples thus have two or more different readings; in other classes such a possibility exists only occasionally.

The semantic interpretation within the third class is 'the object that is no longer in use' (further examples are: *old place*, *flat*, *trail*, *tooth-feeling*, *Band-Aid*, etc.). The most close synonym of *old* here is *former*, and the adjectives *present* and *current* can serve as antonyms in most cases.

The fourth class has the following semantic interpretation: 'the object that has been created in the past and belongs to the past' (further examples are: *old icons*, *silver*, *postcards*; *the old London*, *the old town*, etc.). The most close synonym of *old* here is *ancient*, a possible antonym can be *modern*.

2. We argue that the four interpretation types correlate with the four semantic classes of objects, each of which having its own internal temporal characteristics. Besides, it is interesting to compare these internal temporal characteristics of nouns - in other words, their aspectual characteristics - with the "true" aspectual characteristics of verbs.

Indeed, the first class (old oak) correlates with **gradatives** (the term is borrowed from [Paducheva 1994]) like *increase*, *decrease*, *enlarge*. Gradatives are processes which gradually change the state of objects during a certain period of time. Let us look at the nouns of the first class. What they all denote are natural objects, **gradually** changing in time, such as trees, mountains, etc.

However, the whole picture is a little bit more complicated. The problem is that each language has its own naive concept of what can change in time and what cannot. This does not always correspond to the physical reality. So, according to the Russian naive concepts, not all natural objects do change in time, and, as a result, not all names of objects can combine with *staryj*. For example, sun, sky, fire, stars are conceptualized as stable objects, they don't change and the corresponding nouns do not combine with *staryj*; trees are thought of as changing, but flowers are not, mountains and rocks (but not hills) are considered to be changing, and so are woods (but not steppe). The same opposition holds for animate nouns: so, lions and wolves are capable of gradual change, contrary to worms, flies or, for ex-

ample, nightingales. Monsters and ghosts are also "time-stable", unlike, for example, witches who may well be old.

Since the second class (*old rag*) contains artifacts having a fixed life-time, the natural verbal counterparts for it are **delimitatives** (i.e., verbs denoting an action limited to a certain period of time, cf. Russian *poguljat'* 'to walk for a certain time, to go for a walk'). An interesting consequence of this limited character of the life-time is that all changes affecting the artifacts are considered here as a kind of their "spoiling" (so, an old hat is usually worse than a new one); this effect does not take place in the first class: an old tree is not necessary a bad one, it is only a tree which has undergone a certain change [2].

The third class (*old channel*) resembles verbal **multiplicatives**, i.e. verbs denoting a multiple action consisting of several consecutive acts (cf. Russian *prygat'* 'to jump [several times]', *kashljat'* 'to cough [several times]', etc). Although in these cases each next act repeats the previous one, it does not mean that all the acts within a multiplicative action must be fully identical; moreover, usually they are different. When, for example, one jumps several times, the next jump may well be longer or shorter than the previous one, but from the lexical point of view this is the same situation repeated: the speaker abstracts away from these and similar differences.

If we look back to the nouns, we can state that the nouns of the third class are changeable renewed objects (or parts of objects); each of such objects can (or even must) be changed by a new one, which is absolutely the same from the lexical point of view. Due to this, we can term such objects **cyclic**. A typical example here is a *tyre*: indeed, a tyre is something which is regularly changed after being worn out. Note that the change normally takes place only after the natural life-cycle of the previous object is finished; this is really a very important condition. Indeed, there are cases where we have so-called "permanently renewed" objects (e.g., plates: one changes plates during the dinner not because they are "worn"); one cannot say something like:

? *Put me the dessert in the old plate, please.*

Or imagine the situation when a woman is trying on different coats at the shop: she took two and put them on in turn, and the shop-assistant could advise her:

Oh, take the first one, it is better,

and not:

? *Oh, take the old one, it is better.*

In both situations, the change took place not because the life-cycle of the object was finished, so the conditions for the use of *old* (*staryj*) were not observed.

Finally, the fourth class (*old coins*) could be called a class of **creative** nouns (created at a given moment by a given author). Such objects normally do not decay - they belong to a certain epoch and they reflect it [3]. So, they exist at present as the reflection of the past. As an analogous class of verbs we can take **perfectives** like *remember*, because after the moment of remembering takes place, the result persists for a long time as a long unlimited reflection of the situation.

The fourth class is not large; perhaps, it would be non-productive, if it were not for a new powerful source of these nouns, namely, the words for (technical) devices. The devices are objects which easily become out of date, so, when we say, for example, *old telephone* or *old computer*, we mean that it belongs to the past; the only difference here is that normally the interpretation of the fourth class construction is positive, whereas the subclass of devices yields a negative interpretation (an old computer is rather a bad one, unlike old town or old icon).

3. Now, let us try to implement the results of this classification. We have four types of lexemes, and each of them has its own aspectual characteristics. The important thing is that such characteristics are independent and semantically inherent, because this type of semantic information is indispensable for an adequate semantic description of the lexemes. When explaining what a tree is, we have to mention that trees can grow and change in time, and when explaining what a tyre is, we must mention that tyres are something that is regularly changed after being worn out.

So, if we return to *staryj*, we can (basing on our classification of nouns) describe it as a temporal operator with only one meaning:

- *staryj* = 'that came into being or was created long before the moment of speech'

All the differences in the interpretations of the given phrases with *staryj* can be accounted for the interplay of the meaning of *staryj* and the context of different nouns. So, for the first class *staryj* means 'such that came into being long before the moment of speech [and has changed]'. This can be illustrated

by the following scheme:

For the second class, *staryj* means 'created long before the moment of speech'. Here, 'long before' could be measured, and the measure is the life-time of the object; in other words, the object is so much old that its life-time is nearly finished. This can be illustrated by the following scheme:

The end of life-time is the real reason for the negative judgement in this class; so, it is not a part of the meaning of *staryj*, but a result of the semantic agreement between a noun and an adjective due to the special aspectual characteristics of the given class of nouns.

For the third class, the meaning of *staryj* will be 'such that came into being or was created long before the moment of speech and is no longer in use', because the measure of what is 'long before' here is the life-time of a single referent. So, the object is so much old that it has completed its life-cycle, and the next one has replaced it. Cf. the following scheme:

As for the fourth class, the picture will be similar, only it refers not to the object itself, but to the epoch the object belongs to: the object came into being long ago, so that its epoch has already passed away.

4. Of course, there could be alternative decisions; one of them is dis-

cussed in [Partee 1995]. This paper distinguishes between several semantic types of adjectives: for example, the semantic type exemplified by *skillful* is different from the semantic type exemplified by *former*; every skillful N is an element of a set denoted by N (a skillful surgeon is always a surgeon), but the adjectives of the type *former* do not meet this requirement: a former senator is not a senator. Yet another example is the adjectives of the type *tall*: the values of the dimensional parameter in the contexts like *a tall boy* and *a tall basketball player* are obviously different. So, *tall* forms a different semantic type of adjectives.

These examples show that with such approach all the semantic differences observed are concentrated within the adjective. If we try to apply this approach to *staryj*, the different uses of *staryj* would be distributed in different types of adjectives: the type *former* correspond to our third class (*old channel*), and the type *tall* - to our first class (*old oak*); so, in such a case we shall have three or even four different lexemes *staryj*. On the other side, we have a possibility of an alternative description (presented in this paper) which preserves a single meaning of the adjective despite its broad combinability [4].

Notes

* I would like to express my sincere gratitude to the European Association for Logic, Language and Information (FoLLI) which kindly provided the financial support for my participation in the X-th Amsterdam colloquium.

1 Here I must emphasize that the material of my paper is Russian, so it is not the English word *old*, but the Russian word *staryj* I am concerned with; and all the examples are Russian translated into English, though sometimes it was not easy to do this. It goes without saying that there could be and even must be differences between the linguistic behaviour of Russian *staryj* and English *old*.

2 Curiously, the noun *wine* belongs to the first class and not to the second; it is interpreted not as an artifact, but as a living object. Indeed, *old wine* is closer to *old oak* than to *old hat*, because usually the older the wine is, the better it is.

3 The fourth interpretation is especially favoured by the generic use, because a class or a set of objects is better associated with the epoch than an

individual object.

4 Interestingly, the combinability of the names of kinds of persons (not considered here in detail) confirm this thesis: cf. *old professor* [= 'old-aged'] VS *old friend* [= 'being a friend for a long time'] VS *old president* [= 'former'] VS *old emigration* [= 'of a previous epoch'].

References

Bach, E.: 1968 , "Nouns and noun phrases", in: E.Bach and R.T.Harms (eds), *Universals in linguistic theory*, Holt, N.Y.

Dahl, Ö.: 1970 , "On points of reference", *Semantikos, 1.1.*

Givón, T.: 1979 , *On understanding grammar*, Academic press, N.Y.

Jakovleva, E.: 1994 , *Fragmenty russkoj jazykovoj kartiny mira* (Fragments of Russian linguistic world image), Gnozis, Moscow.

Paducheva, E.V.: 1994 , "Taksonomicheskie kategorii glagola i semantika vidovogo protivopostavlenija" (Taxonomic categories of verbs and semantics of aspectual oppositions), *Semiotika i informatika*, 34.

Partee, B.H.: 1995 , "Lexical semantics and compositionality", in: L.Gleitman, M.Liberman (eds), *Invitation to cognitive science, part I: Language*, MIT, Cambridge (MA).

Vendler, Z.: 1967 , *Linguistics in Philosophy*, Cornell University Press, Ithaca.

An event-based semantics for deontic utterances

Gene Rohrbaugh

The University of Texas at Austin

Abstract

Several problems for modeling the semantics of deontic utterances are linked to the choice of possible worlds as theoretical primitives. Possible worlds are shown to be responsible for three problems for deontic semantics: Lewis's *problem about permission,* von Wright's *paradox of free choice permission,* and a problem for permissions involving a certain type of quantification. An alternative formulation of the semantics of deontic utterances is proposed, in a framework using states as primitives. Events, which are defined in terms of states, are used to represent the changes brought about by deontic utterances.

1. Introduction: deontic modeling

A well known and productive analysis of the illocutionary force of deontic utterances, specifically commands and permissions, can be traced back to David Lewis in 1970. Though the relevant paper remained unpublished until 1979, Lewis's ideas are already apparent in Kamp's (1973) "Free choice permission" and his later (1978) "Semantics versus Pragmatics." At the heart of Lewis's analysis is the observation that commands restrict the options of the addressee, while permissions broaden them. Much of Lewis's and Kamp's further discussion centers on how the options of an individual are changed through commands and permissions.

Underlying the work of both Lewis and Kamp is the possible worlds semantic framework. As in Montegovian possible world semantics, the meaning of a proposition is represented as a set of possible worlds — those worlds in which the proposition is true. Furthermore, Lewis proposes using a set of possible worlds to represent the options of an individual — those worlds which he or she is authorized to bring about. This allows us to separate the illocutionary force of commands and permissions from their propositional content. The restriction brought about by a command can be specified by taking the intersection of the prior option set with the propositional content of the command. For example, when the command in (1) is uttered[1], Joseph's option set contracts to include only those worlds in which he eats his lima beans.

(1) Joseph must eat his lima beans.

Similarly, the effect of a permission can be reflected by set union, adding to the addressee's option set worlds in which the propositional content of

[1] under appropriate circumstances by a person in authority.

the permission is true. Thus the permission in (2) adds to Joseph's option set (at least some) world(s) in which he eats a cookie.

(2) Joseph may eat a cookie.

In this paper, I will argue for maintaining Lewis's initial observation about the illocutionary function of commands and permissions, while using events instead of possible worlds to represent both the propositional content of deontic utterances and the options of the conversational participants. This will, I believe, provide at least a partial solution to several problems for deontic semantics, including those discussed by Lewis and Kamp, along with another problem discussed in Rohrbaugh (1995).

2. Problems for possible world modeling

This section will examine three problems that arise for the Lewis/Kamp (hereafter L/K) analysis within the possible worlds framework: the spurious permission problem (Lewis's problem about permission), the free choice permission problem (von Wright's paradox of free choice permission, as discussed by Kamp), and the quantified permission problem (Rohrbaugh, 1995).

2.1 The spurious permission problem

Under Lewis's analysis, the change brought about by a command can be described as *complete,* in that it affects all worlds in the propositional meaning. The command "Joseph must weed the garden," removes from Joseph's option set every world in which he does not weed the garden. In contrast, the change brought about by a permission is incomplete. The permission "Joseph may eat a cookie" does not add to Joseph's option set every world in which he eats a cookie, as some of these may be worlds in which he also happens to beat the cat or break a lamp. These worlds should not be added to Joseph's option set when Mother says "Joseph may eat a cookie."[2] Such a change would be spurious, going beyond the intended effect of the permission. It seems that we should specify some smaller change to the addressee's option set than the maximal one, and therein lies the spurious permission problem.

Consider again the permission in (2) along with the model in Figure 1, where possible worlds are represented as histories.

[2] assuming the latter are impermissible.

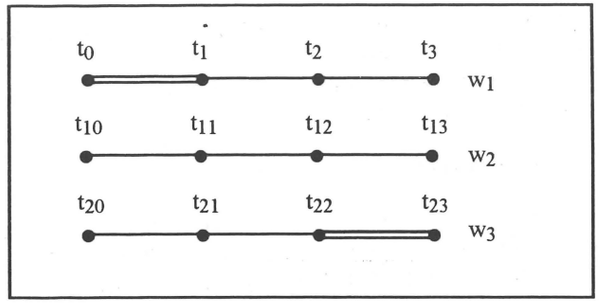


Figure 1

Three possible worlds are represented, each a succession of time points. Segments with double lines represent portions of histories in which *Joseph eats a cookie.* If the effect of a permission were complete, "Joseph may eat a cookie" would add w_1 and w_3 to Joseph's option set, as these are the worlds in which the propositional content of the permission is true. The problem with this is that we have said nothing about what else happens in these worlds. In w_1, the segment between t_2 and t_3 might represent an event of *Joseph-beating-the-cat.* In that case w_1 should be excluded and only w_3 added to Joseph's option set. For Lewis, the spurious permission problem is the problem of determining which worlds to add to the option set.

Lewis discusses five possible solutions, each focused on finding some subset of the propositional content and adding it to the option set of the addressee. Perhaps the most promising is one based on a suggestion by Robert Stalnaker which relies on reprehensibility gradings to determine which world(s) are to be added. It seems reasonable to say that a world in which Joseph eats a cookie and beats the cat is more reprehensible than one in which he, for example, eats a cookie and plays quietly in his room. We could say then that a permission PER(ϕ) adds the least reprehensible ϕ world(s) to the option set of the addressee.[3] For the model described above, this would amount to adding w_3 and omitting w_1, as the former is less reprehensible than the latter. This is certainly better than the *complete* solution (adding both worlds), but is it enough? We still have something very much like the spurious permission problem. We have added not only the *Joseph-eats-a-cookie* event in $<t_{22}, t_{23}>$, but also those events represented by $<t_{20}, t_{21}>$ and $<t_{21}, t_{22}>$. What are these events, and why should permission to *eat-a-cookie* carry them along?

[3] where PER represents the permission illocutionary operator and ϕ the propositional content.

It seems we might profit by recasting Lewis's analysis against a framework with a way to refer to elements more fine-grained than possible worlds to represent more precisely the effect of a permission. This is what I will attempt in Section 3.

2.2 The free choice permission problem

Kamp (1973) discusses free choice permission within the same possible worlds framework as that used by Lewis. The free choice permission problem can be illustrated using the examples in (3):

(3) a. Joseph may eat an apple.
b. Joseph may eat an apple or a pear.

Intuitively the permission in (b) entails that in (a). The problem is that most formal systems of logic predict the reverse, that (a) entails (b), based on the law of disjunction introduction.[4]

In order to solve this problem, Kamp (1978) proposes that the meaning of a disjoined permission PER($\phi \vee \psi$) is derived first by computing PER(ϕ) and PER(ψ) independently, and then taking the union of the two as the 'meaning' of PER($\phi \vee \psi$). Thus Kamp interprets *or* as disjunction at the speech act level. This approach is problematic for two reasons. First, the proposed equivalence fails in cases such as in (4):

(4) a. I permit you to eat an apple or an orange.
b. ≠ I permit you to eat an apple or I permit you to eat an orange.

If the interpretation of *or* is to be represented uniformly by set union (what Merin (1992:108) calls *uniform boolean recursion — UBR*), then certainly the interpretation of (b) should follow the course mapped by Kamp, and yet it is (a) and not (b) that has a free-choice reading. At best (b) is used mischievously to say something like "I'm willing to enact one of these permissions, but I won't tell you which one."

Even more serious, performative versions of permissions cannot be disjoined at all, as illustrated in (5).

(5) a. I hereby permit you to eat an apple or a pear.
b. #I hereby permit you to eat an apple or I hereby permit you to eat a pear.

In this case (a) is an example of free choice permission, granting the addressee the freedom to choose either option. Yet the permission in (b), presumably much closer in surface form to Kamp's solution, not only fails

[4] That is, for propositions p and q, p entails (p ∨ q). E.g., the truth of *Joseph ate an apple* is sufficient to ensure the truth of *Joseph ate an apple or a pear.*

to have the same meaning as (a), it cannot be used at all.[5] It seems desirable then to find another solution to the free choice permission problem which maintains UBR while allowing the disjunction operator to apply at the propositional level (or at least at a sub-speech act level).

2.3 The quantified permission problem

In the previous section we examined the free choice permission problem, which could be summed up by saying that our formal systems of logic predict the wrong entailments between simple and disjoined permissions. A similar problem arises for quantified permissions. Consider the example in (6):

(6) a. Joseph may eat three cookies.
b. Joseph may eat a cookie.

Standard intuitions dictate that (a) entails (b). In the absence of the *exactly* reading of the quantifier, the addressee of (a) could cooperatively respond by eating one cookie, or two cookies, or three cookies. In contrast the addressee of (b) could not (cooperatively) respond by eating three cookies, hence the intuition that (a) entails (b) and not vice versa. Alternately we could say that the quantifier position in the permission in (a) is usually given an *at most* reading rather than *at least* reading. The possible worlds framework not only fails to predict this, it predicts the reverse. To see why this is so, consider the case given. According to the possible worlds framework, the permission in (a) should add to Joseph's option set those worlds in which he eats three cookies. In which worlds is it true that he eats three cookies? Of course it is true in worlds where he eats exactly three cookies, but it is also true in worlds where he eats four cookies, or five cookies, or six cookies, etc. It is not true in a world where he eats only two cookies or one cookie. Hence the possible worlds framework wrongly predicts an *at least* reading for quantified permissions such as (6).

Note that the possible worlds framework does not have a problem predicting the *exactly* reading for quantified permissions, as in that case there is no entailment relationship between the permissions in (6). But given that the *at most* reading is available, and perhaps even more salient, we should find a way to make the correct prediction about the direction of entailment for quantified permissions.

[5] I think this is because the use of the performative version conveys that the speaker intends to have a certain perlocutionary effect. In contrast, the disjunction conveys that the speaker either does not know or does not care which effect his utterance will have. These two meanings conflict in a example such as (5.b)

3. Solutions in event modeling

The event-based framework I will examine in this paper could be seen as a recasting of the L/K apparatus for deontic modeling in a framework incorporating Dowty's (1979) branching time structures. The basic intuition of Lewis remains in that individuals are assumed to have associated option sets, and permissions have the effect of widening the option set of the addressee. This widening is determined by the propositional content of the permission. The major difference is that now events, rather than possible worlds, are the elements that make up both option sets and propositional meanings.

Let us begin with Dowty's observation that possible worlds can be represented as linear histories. We need the following definitions:

S: a set of states of affairs s, s', s'', etc., ordered by a temporal precedence relation $<_t$. Read $s <_t s'$ as "s temporally precedes s'."

$<_t$ must be irreflexive, transitive, and asymmetric (that is, a strong order), and in addition backwards linear. Backwards linearity is defined as follows:

$\forall$ s, s', s'' [[$s <_t s'' \wedge s' <_t s''$] $\rightarrow$ [$s <_t s' \vee s' <_t s$]]

This condition basically ensures that histories may branch apart moving into the future, but they can never rejoin.

My terminology here differs from Dowty's , as he begins with T, a set of similarly ordered time points t, t', t'', etc. The difference is only terminological, as his time points represent worlds-at-a-time. A world/time pair can be associated with the state of affairs that holds in that world at that time, and hence the two systems are effectively the same.

Given Dowty's system we now have that states-of-affairs are primitively defined, and possible worlds are constructed from sets of states-of-affairs.[6] Thus we need not lose any of the power of possible worlds modeling, but we have in addition a notion more fine-grained to work with.

Stative propositions such as those in (7.a) are easily defined in this model as the set of states in which the proposition holds.

(7) a. The grass is wet.
b. The grass got wet.

[6] specifically, worlds are maximal chains in a branching time structure.

The case is not quite so straightforward for eventive propositions such as (7.b). We cannot represent the meaning of an eventive proposition ϕ as merely the set of states in which ϕ holds, as eventive propositions do not hold at a state; rather they seem to entail state-change. It seems reasonable then to define eventive propositions in terms of state-changes. Consider the branching time structure in Figure 2.

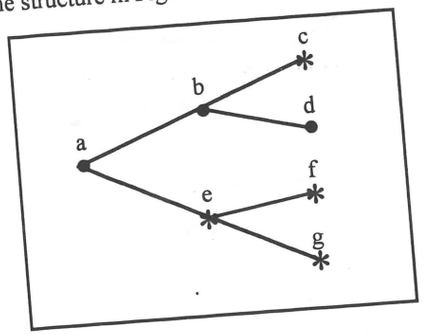


Figure 2

The states marked with asterisks are those in which the grass is wet, those marked with simple dots are those in which it is not wet. Within this model the two propositions in (7) can be represented as follows:

(8) a. [[the grass is wet]] = {c, e, f, g}
b. [[the grass got wet]] = {<b, c>, <a, e>}

Why not include <a, f> and <a, g> in the meaning of the (b) as well? These are also changes from *dry-grass* states to *wet-grass* states. Suffice it to say that I assume a certain granularity level with which we view the world. Events in this system are represented as transitions which are in some sense minimal.

We can now recast the L/K apparatus for deontic utterances in terms of events rather than possible worlds. The option set is now a set of events, those events which are permissible, and the effect of a permission is to add to the option set the propositional content of the permission. Consider again the permission in (2) above, and its predicted effect given the model in Figure 3.

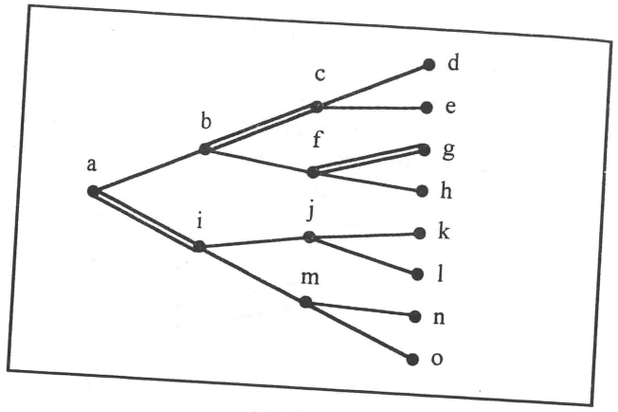


Figure 3

Segments marked with a double line indicate events of Joseph eating a cookie. The effect of the permission in (2) is to add {<b, c>, <f, g>, <a, i>} to Joseph's option set. In general the effect of a permission can be defined as follows:

Definition 1: the PER operator

$PER_{\alpha}(\phi)$: $OPT_{\alpha} \rightarrow OPT_{\alpha} \cup [\![\phi]\!]$

Note that the permission illocutionary operator is intended to apply to eventive propositions; if the option set is indeed a set of events, we should not be able to add simple states to it. This is not an unwelcome result, as it may contribute to an explanation of why eventive propositions seem to be more at home in commands and permissions than stative ones, as illustrated in (9).

(9)
a. You may eat an apple.
b. You may stand tall.
c. Get rich quick!
d. #You may be eating an apple.
e. #You may be tall.
f. #Be rich now!

We might explain the acceptability of (a) - (c) as a consequence of the fact that their propositional contents are represented as sets of events, whereas those in (d) - (f) are represented as sets of states. Then the appropriate deontic illocutionary operators should select eventive propositions and disallow stative ones within their scope.[7] Within the possible worlds

[7] Obviously things are not quite so simple, as some statives can occur in deontic utterances. Consider the following:

framework this would not be possible, as all of the propositional contents would be represented in the same way, as sets of possible worlds.

3.1 Towards a solution for spurious permission

To see the improvement over the possible worlds approach, consider the following model, where *Joseph-eats-a-cookie* events are marked with double lines and *Joseph-beats-the-cat* events are marked with dashed lines.

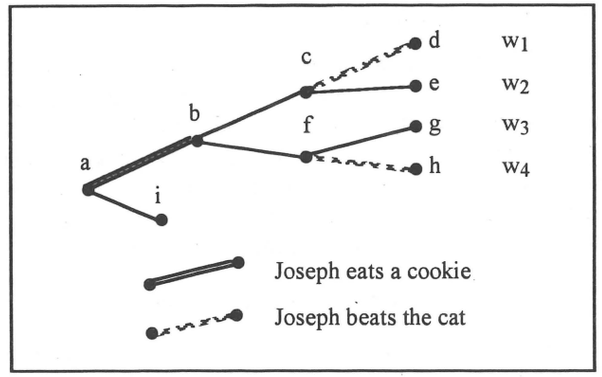


Figure 4

The possible worlds approach would look for the set of worlds in which Joseph-eats-a-cookie is true, in this case $\{w_1, w_2, w_3, w_4\}$. Prior attempts to solve the spurious permission problem would then focus on finding an appropriate subset of this set to add to Joseph's option set. The event-based semantics merely says to add the event represented by segment <a, b> to the option set, and thus the permission has no effect on events which are not *Joseph-eats-a-cookie* events (specifically the later *Joseph-beats-the-cat* events <c, d> and <f, h>).

One might wonder if this system does in fact make different predictions than Lewis's account. After all, we do still have possible worlds, even though they are constructed rather than primitive. Would the two systems make different predictions about which worlds are sanctioned? In Lewis's system the prediction is clear: a world is sanctioned if it is in the

i) Remain calm!
ii) Don't be anxious!
iii) Don't worry; be happy!

It may be that such utterances are attempts to require or to proscribe events that lead to the states referred to, rather than the states themselves. E.g., in order to 'be happy', one should think happy thoughts, take a brisk walk, get plenty of rest, etc.

option set. For each world in $\{w_1, w_2, w_3, w_4\}$ we either add the world or not. If we add any one of these, the entire world becomes sanctioned. If we do not add one world, the permission has no effect.

In the event-based system, the option set consists of events rather than worlds, so what would it mean for a world to be sanctioned? I would like to define it as follows:

Definition 2: Sanctioning of worlds

A world is sanctioned if and only if it consists entirely of transitions which are in the option set.

Given this definition, we see that a permission can have an effect without sanctioning any worlds. Given the model above in Figure 4, the permission in (2) would add only <a, b> to the option set. No world would (necessarily) be sanctioned by this move. Later (or earlier) permissions might add (for example) <b, c> and <c, d> to the option set, rendering w_1 a sanctioned world, but this need not happen. In general, the sanctioning of worlds arises due to the cumulative effect of permissions.

3.2 Towards a solution for free choice permission

The solution I will give for the free choice permission problem is actually quite close to Kamp's earlier (1973) proposal, which he later abandons in favor of the solution sketched in Section 2.2 above. Before offering my solution then, I should present Kamp's proposal and his motivation for abandoning it.

In order to refer to the entailment-like relation between the permissions in (3), repeated here, Kamp (1973) introduces a notion of p(ermission)-entailment.

(3) a. Joseph may eat an apple.
b. Joseph may eat an apple or a pear.

Definition 3: Kamp's (1973) p-entailment

... for arbitrary sentences *s* and *s'*: *s p-entails s'* iff in any situation in which both *s* and *s'* can be used to grant permissions the class of actions which would be moved by a permission statement involving *s* would include the class of actions moved by a permission statement involving *s'*.[8]

According to Kamp's earlier 1973 analysis, the logical connective *or* applies at the propositional level, rather than at the speech act level as in his

[8] Kamp 1973:66.

1978 analysis. (b) moves into the option set those worlds which are in ⟦Joseph eats an apple or a pear⟧, simply represented as ⟦Joseph eats an apple⟧ ∪ ⟦Joseph eats a pear⟧. Thus (b) p-entails (a) because it moves into the option set all of the worlds added by (a) alone.

Kamp's motivation to abandon this solution seems to be linked to the lack of a solution for the spurious permission problem. Perhaps the most promising solution available at the time was the Lewis/Stalnaker account using reprehensibility gradings. This gives the right result under the assumption that *eating-an-apple* and *eating-a-pear* are equally reprehensible. (b) adds the least reprehensible *eating-an-apple* worlds and the least reprehensible *eating-a-pear* worlds. A problem arises, however, in a situation where eating an apple is much more reprehensible than eating a pear, for example because pears grow in the garden, while apples must be bought at the store. In such a circumstance the Lewis/Stalnaker approach predicts that (b) will add only *eat-a-pear* worlds, as these are the least reprehensible *eat-an-apple-or-a-pear* worlds. This prediction conflicts with the intuition that (b) moves both *eat-an-apple* worlds and *eat-a-pear* worlds into permissibility.

With the new event-based system, we no longer rely on reprehensibility gradings to solve the spurious permission problem, and hence Kamp's early solution to the free choice permission problem might be considered once again. Maintaining Definition 1, the derivation of a disjoined permission is straightforward.

$$\text{PER}_\alpha(\phi \vee \psi)\text{: } \text{OPT}_\alpha \rightarrow \text{OPT}_\alpha \cup [\![\phi \vee \psi]\!]$$
$$= \quad \text{OPT}_\alpha \cup ([\![\phi]\!] \cup [\![\psi]\!])$$

Note that UBR is maintained, with *or* represented as union. The free choice entailment is easily explained as the derivation of a disjoined permission moves into the option set all of the events which would be moved by a simple permission. And finally, the problem for the Lewis/Stalnaker account does not even arise, as we make no reference to reprehensibility gradings.

3.3 Towards a solution for quantified permission

In this section I will only sketch a solution for the quantified permission problem, and even this sketchy account relies on a part relation on events which I cannot yet formally define. Intuitively what we want to say is that whenever ψ is a natural part of ϕ, a permission to ϕ 'includes' permission to ψ. For example, permission to eat an apple normally seems to include permission to bite the apple, to chew the apple, to swallow the

apple, etc., but not include permission to throw the core at a passerby, even if one somehow works this into his apple-eating routine. This seems to be the effect that Kratzer (1989) describes and terms *lumping*. In her exposition, Kratzer uses the term *lumping* to refer to the relationship between painting a still life and painting an apple that is part of the still life. She says that in a given situation, the proposition "Paula painted a still life" *lumps* the proposition "Paula painted an apple," but not the proposition "Paula drank some tea," even though the *tea-drinking* may have spatio-temporally overlapped with the *still-life-painting*. Similarly, for our purposes we might say that eating an apple *lumps* biting it, chewing it, swallowing it, etc., but not throwing the core at a passerby. The relation we are after is certainly weaker than entailment, as not every event of *Joseph-eats-an-apple* contains a sub-event of *Joseph-chews-an-apple* (he might have pureed it first). Nevertheless, our intuitions suggest that biting an apple is a normal step in eating it.

Despite the apparent similarities, Kratzerian lumping is not what I need for my natural part relation. Kratzer defines lumping within the framework of situation semantics. She demonstrates convincingly that lumping cannot be defined in terms of spatio-temporal slices of possible worlds. Thus we cannot define lumping in our current system, as branching time structures merely give us spatio-temporal slices. Neither can we use Kratzer's version of situation semantics to solve our problems. For Kratzer, the meaning of a proposition ϕ is the set of (all) situations in which ϕ is true. If we attempt to specify the change induced by a permission in terms of situations, the spurious permission problem returns, as the set of situations in which Joseph eats a cookie may include situations in which he also beats the cat. We are forced again to find some subset of the propositional content to add to the option set[9], but it is not clear which subset to choose.

Assuming a natural part relationship, represented by $\sqsubseteq_n$, we can redefine the effect of a permission as follows:

Definition 4: PER in terms of *natural parts*

$PER_\alpha(\phi)$: $OPT_\alpha \rightarrow OPT_\alpha \cup [\![\psi]\!]$, for all ψ such that $\psi \sqsubseteq_n \phi$.

Definition 4 predicts that permission to eat three cookies would include permission to eat one cookie, as an event of eating three cookies might naturally contain a sub-event of eating one cookie. The simple solution has a simple problem, though, in that it requires all events to be represented as simple. In determining if Joseph is within his rights eating two

[9] which would in this case be represented as a set of situations.

cookies after having been granted permission to eat three, we can conclude that it is permissible only if we represent eating-two-cookies as a single event. As it is, the model can only add to the option set single uninterrupted events, represented by minimal state changes. But it seems desirable to reflect the permissibility, under such a permission, of Joseph eating one cookie, then straightening up his room, and then eating a second cookie. What is needed seems to be a way to allow the quantification of the permission to be a quantification over events. In essence, the permission to eat three cookies adds to the option set any cookie-eating event which it is not preceded by three cookie-eating events already. Consider the permission in (10).

(10) Joseph may eat two cookies.

OPT should be increased by adding any event of *Joseph-eats-a-cookie* provided that it is not preceded by two other *Joseph-eats-a-cookie* events. In the case of the model in Figure 5, this means that {<a, b>, <c, d>, <g, h>, <b, j>} are added, while {<d, e>, <j, k>} are not.

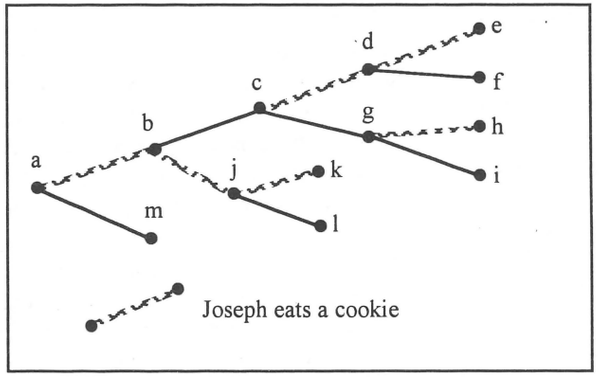


Figure 5

The natural part relation could also play a role in the solution to the full choice conjunction problem[10]. Consider the permission in (11).

(11) Joseph may have cake and ice cream.

There is a reading, the so-called *package deal* reading, which says that Joseph could take up one of the disjuncts only if he took up both. A more natural reading is the one which allows Joseph to have both, either, or nei-

[10] see Rohrbaugh 1995 for further discussion of full choice conjunction.

ther. Definition 4 predicts this reading since *having-cake* is a natural part of *having-cake-and-ice-cream.*

So we have seen that by manipulating events rather than possible worlds we have been able to reduce some of the problems for deontic semantics.

4. Further directions

I see at least two problems for further exploration within this framework. First, the implied discreteness of events does not allow for disjoint fulfillment of a permission.[11] If Joseph is granted permission to eat a cookie, normally we would not require that the cookie-eating be a single continuous event. He might eat half of it, then watch some TV, eat another quarter, walk to the mailbox, etc. The solution to this might lie along the same path as that proposed for the quantified permission problem. A permission to ϕ should add to the option set all events that are a natural part of a ϕ event.

Second, I have not addressed the problem of overlapping events in this paper.[12] Events are represented as pairs of consecutive states. It is seldom, if ever, the case that only one aspect of the state of affairs changes from one state to the next. If Joseph eats an apple between states *a* and *b,* it is likely that other things have changed as well. It may have started raining, lightning may have struck, or Joseph may have broken a lamp, during the same interval. We would like the permission "Joseph may eat an apple" to avoid adding those apple-eating events which also happen to be lamp-breaking events, but as with possible worlds there is no general way to distinguish. Thus we have something like the spurious permission problem all over again. We have rightly avoided adding events which occur within the same history as a permitted event at a different time, but we have failed to exclude events which temporally overlap the permitted event completely. For example consider the model in Figure 6.

[11] As for the discreteness itself, I imagine some might argue that branching time structures are not discrete, and the event-based system is thus unrealistically artificial. I would only say that the discreteness in the modeling system need not reflect discreteness in the real world; it might arise due to a certain granularity level with which we view the world. Certainly there are sub-events to an event such as eating an apple, but for the purposes of conversation we might act as if the apple-eating were condensed into a single discrete event.

[12] See Rohrbaugh (1995) for a more complete discussion of overlapping events.

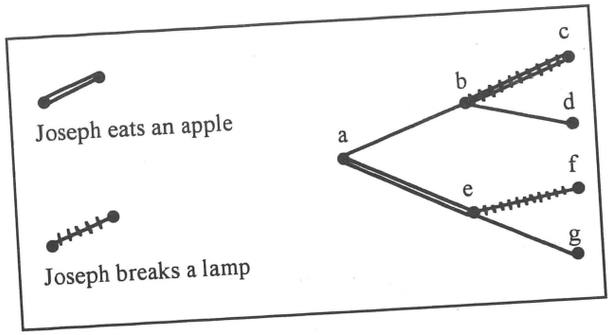


Figure 6

The possible worlds approach would add three worlds to Joseph's option set, those determined by the histories <a, b, c>, <a, e, f>, and <a, e, g>, as each of these is a world in which Joseph eats an apple. Of these, the first two introduce spurious permission effects, allowing the lamp-breakings in <b, c> and <e, f>. The event-based approach correctly omits the <e, f> *lamp-breaking* event, but still wrongly includes that in <b, c>.

Nick Asher has suggested that one possible solution would be to further subdivide events, not temporally but qualitatively. We could refer to event tokens not as simple pairs of states, but as subscripted pairs of states. Recasting the model in Figure 6 we would have something like Figure 7.

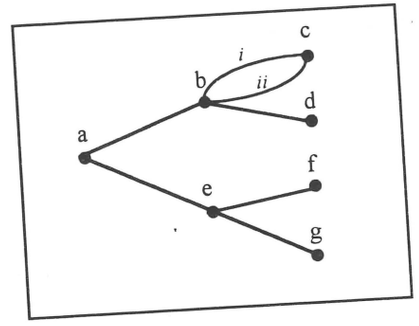


Figure 7

The relevant token event of apple eating could be referred to as $\langle b, c\rangle_i$ and the temporally overlapping token event of lamp breaking as $\langle b, c\rangle_{ii}$. Then we can distinguish once again the relevant event, saying that the permission to eat an apple would add $\{\langle b, c\rangle_i, \langle a, e\rangle\}$ to the option set.

5. References

Kamp, Hans: 1973, Free Choice Permission, in *Proceedings of the Aristotelian Society* Volume LXXIV, 57-74. London: Methuen & Co. Ltd.

Kamp, Hans: 1978, Semantics Versus Pragmatics, in Guenthner, F. and Schmidt, S. J. (editors), *Formal Semantics and Pragmatics for Natural Language*, 255-287. D. Reidel Publishing Co., Dordrecht.

Kratzer, Angelika: 1989, An Investigation of the Lumps of Thought, *Linguistics and Philosophy* 12(5), 607-653.

Lewis, David: 1979 A Problem about Permission, in Saarinin, E., Hilpinen, R., Niiniluoto, I., and M. Provence Hintikka (editors), *Essays in Honour of Jaakko Hintikka*, 163-175. Dordrecht: D. Reidel Publishing Company.

Merin, Arthur: 1992, Permission Sentences Stand in the Way of Boolean and Other Lattice-Theoretic Semantices, *Journal of Semantics* 9(2), 95-162. Oxford University Press.

Rohrbaugh, Eugene: 1995, *Scalar Interpretation in Deontic Speech Acts*, Ph.D. dissertation, University of Texas, Austin.

von Wright, G. H., An Essay on Deontic Logic and the Theory of Action, Amsterdam 1969.

On Modes of Operation*

Anna Szabolcsi
UCLA

1. Overview

In traditional syntactic and semantics theories, the behavior of all noun phrases is presumed to be uniform. This is expressed by the fact that the rules used to account for noun phrase behavior do not distinguish between types of noun phrases.

The first theory that dramatically rejects this assumption is DRT; it primarily uses evidence pertaining to pronominal anaphora.

Of course, anteceding pronouns is not the only thing noun phrases do. For instance, they take scope. Again, the traditional assumption is that noun phrases are alike in this respect (or, if they are not, the differences do not merit serious investigation). But this is wrong. As anecdotal (but serious) evidence, consider what happens when you teach quantifier scope ambiguities to an unbrainwashed undergraduate class. It takes long hours to come up with a few examples that the students accept on the readings the standard theory predicts. This does not speak well of the theory.[1]

---

* This text is more or less identical to the one read at the colloquium. A more elaborate version is "Strategies for Scope Taking" in Szabolcsi (1996).

[1] More concretely, note the following small sample of data, an account of which can be found in Stowell and Beghelli (1994, 1995). Only (i) has a natural object wide scope reading that traditional theories make us expect:

(i) Two men read every book.
✓ 'for every book, there is a possibly different set of two men who read it'

In (ii), the plural indefinite object may be referentially independent of the subject but can nevertheless not make the subject referentially dependent:

(ii) Two men read three books.
* 'for each of three books, there is a possibly different set of two men who read it'
✓ 'there are three books such that (the same) two men read each'

In (iii), the modified numeral object can barely scope over another modified numeral in subject, and in (iv), it cannot scope over a universal subject at all.

(iii) At most two men read more than five books.
?? 'for each of more than five books, there is a possibly differet set of at most two men who read it'

(iv) Every man read more than five books.
* 'for each of more than five books, every man read it'

5. References

Kamp, Hans: 1973, Free Choice Permission, in *Proceedings of the Aristotelian Society* Volume LXXIV, 57-74. London: Methuen & Co. Ltd.

Kamp, Hans: 1978, Semantics Versus Pragmatics, in Guenthner, F. and Schmidt, S. J. (editors), *Formal Semantics and Pragmatics for Natural Language*, 255-287. D. Reidel Publishing Co., Dordrecht.

Kratzer, Angelika: 1989, An Investigation of the Lumps of Thought, *Linguistics and Philosophy* 12(5), 607-653.

Lewis, David: 1979 A Problem about Permission, in Saarinin, E., Hilpinen, R., Niiniluoto, I., and M. Provence Hintikka (editors), *Essays in Honour of Jaakko Hintikka*, 163-175. Dordrecht: D. Reidel Publishing Company.

Merin, Arthur: 1992, Permission Sentences Stand in the Way of Boolean and Other Lattice-Theoretic Semantices, *Journal of Semantics* 9(2), 95-162. Oxford University Press.

Rohrbaugh, Eugene: 1995, *Scalar Interpretation in Deontic Speech Acts*, Ph.D. dissertation, University of Texas, Austin.

von Wright, G. H., An Essay on Deontic Logic and the Theory of Action, Amsterdam 1969.

On Modes of Operation*

Anna Szabolcsi

UCLA

1. Overview

In traditional syntactic and semantics theories, the behavior of all noun phrases is presumed to be uniform. This is expressed by the fact that the rules used to account for noun phrase behavior do not distinguish between types of noun phrases.

The first theory that dramatically rejects this assumption is DRT; it primarily uses evidence pertaining to pronominal anaphora.

Of course, anteceding pronouns is not the only thing noun phrases do. For instance, they take scope. Again, the traditional assumption is that noun phrases are alike in this respect (or, if they are not, the differences do not merit serious investigation). But this is wrong. As anecdotal (but serious) evidence, consider what happens when you teach quantifier scope ambiguities to an unbrainwashed undergraduate class. It takes long hours to come up with a few examples that the students accept on the readings the standard theory predicts. This does not speak well of the theory.[1]

---

* This text is more or less identical to the one read at the colloquium. A more elaborate version is "Strategies for Scope Taking" in Szabolcsi (1996).

[1] More concretely, note the following small sample of data, an account of which can be found in Stowell and Beghelli (1994, 1995). Only (i) has a natural object wide scope reading that traditional theories make us expect:

(i) Two men read every book.
✓ 'for every book, there is a possibly different set of two men who read it'

In (ii), the plural indefinite object may be referentially independent of the subject but can nevertheless not make the subject referentially dependent:

(ii) Two men read three books.
* 'for each of three books, there is a possibly different set of two men who read it'
✓ 'there are three books such that (the same) two men read each'

In (iii), the modified numeral object can barely scope over another modified numeral in subject, and in (iv), it cannot scope over a universal subject at all.

(iii) At most two men read more than five books.
?? 'for each of more than five books, there is a possibly differet set of at most two men who read it'

(iv) Every man read more than five books.
* 'for each of more than five books, every man read it'

The assumptions of DRT have some consequences for scope; e.g. the noun phrases that have discourse referents corresponding to them are predicted to scope quite freely. This is on the right track but does not account for the observed scopal diversity in its entirey.

The present paper is part of a larger joint enterprise that has looked at the diverse behavior of quantifiers, with special emphasis on scope. The results are coming out in a volume entitled "Ways of Scope Taking."

In this talk I will begin by looking at the expression of scope in a language that has come to be known to "wear its logical form" on its sleeve: Hungarian. The critical question is, how does Hungarian disambiguate scope in its surface syntax? Does it simply line up the quantifiers in the desired scopal order?

The answer is No. Work on Hungarian over the past 15 years or so has shown that Hungarian quantifiers occur in specific, syntactically distinguishable positions. They are reviewed in Section 2. Since most quantifiers have more than one specific position to occupy, this strategy eventually yields a string in which, by and large, the wider scope quantifier precedes the narrower scope one. However, the important point is that this is not achieved by simply creating a scopal sequence, but by letting the quantifiers enter into an articulated structure.

That is to say, the Hungarian data show that SCOPE LIVES OFF OF AN INDEPENDENT STRUCTURE.

I may immediately note that this observation is not specific for Hungarian. Stowell and Beghelli (1994, 1995) have argued that the Logical Form of English sentences has the same property -- and it turns out that the positions they postulate match up quite straightforwardly with those that Hungarian has in its surface syntax. Thus, although in this talk I will be arguing on the basis of Hungarian (and the Hungarian data are sometimes critical in devising the argument), the overall claims are assumed to hold also for English.

So, what is this structure that scope lives off of? I will argue that it is a structure whose positions correspond to two basic modes of operation for noun phrase denotations. These modes of operation correspond to canonical verification procedures.

One, a noun phrase may act as a logical subject of predication. I.e. its denotation may provide an entity that the (corresponding) predicate of the sentence is "about." I will call this the predicational mode. Two, a noun phrase may act as a classical generalized quantifier, specifically, as a counting quantifier, operating on the denotation of a predicate. I will call this the counting mode, or predicate operator mode.

Are these modes of operation related to anything familiar? Yes. I

will suggest that the two modes correspond, although conceptually are not identical to, the basic DRT distinction between introducing a discourse referent and acting as a generalized quantifier (in Kamp and Reyle's 1993 terminology). This claim amounts to adding a procedural flavor to DRT. The suggestion is that DRT need not (should not?) be simply regarded as a more elaborate logical syntax; for instance, discourse referents are not simply variables to be closed existentially. Rather, introducing a discourse referent amounts to producing a logical subject of predication, an entity that contributes to interpretation in a particular manner.

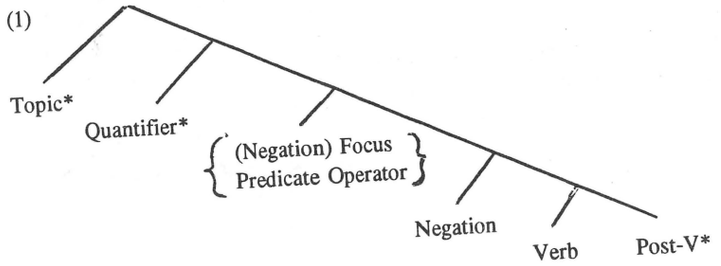
Apart from this conceptual revision, does this proposal entail any concrete modifications of the standard DRT assumptions? Yes. The most important of these pertains to universals like *every dog*. I will argue that they also operate in the predicational mode and, in fact, introduce a discourse referent of some sort. Hungarian has a several non-universal quantifiers that pattern the same way, and we shall see that their logical properties are critical in developing this argument.

Finally, how do these modes of operation relate to denotational semantic properties? I will show that some properties, specifically monotonicity properties, crucially delimit the options a quantifier has; however, they do not determine its options. That is, it can be shown that quantifiers that to my best knowledge are denotationally equivalent may well occur in distinct surface syntactic positions in Hungarian and, accordingly, exhibit distinct modes of operation.

2. The basic facts of Hungarian

So, let us look at Hungarian. The data have a generative syntactic analysis, but since nothing here turns on this, in what follows I will use the most theory neutral descriptive names for the positions that exist in the Hungarian literature. This structure was established in the early Eighties by László Hunyadi, Katalin Kiss, and myself, among others.

(1)



The above positions can be distinguished by surface syntactic tests. E.g. a Topic, but not a Quantifier, can be followed by sentential adverbials like tegnap 'yesterday;' if the verb has a separable prefix, it is is preverbal position if the verb is directly preceded by a Quantifier but occurs postverbally if the verb is directly preceded by a Focus or a Predicate Operator. These details are immaterial for present purposes, all you have to know is that one does not need to belong to any specific school of theoretical syntax to recognize the distinctness of the above syntactic positions.

In this paper, I will be focusing on just three positions: Topic, Quantifier, and Predicate Operator, and ignore the rest.

These positions can be filled simultaneously and, as was mentioned above, their linear order determines the scopal order of the quantifiers (this latter fact is not reflected in the idiomatic translation). E.g.,

(2) [$_{\text{Topic}}$ A tanár [$_{\text{Quant}}$ minden gyereknek [$_{\text{Quant}}$ számos példát
the teacher every child-to several problems-acc
[$_{\text{PredOp}}$ ötnél többször [$_{\text{Verb}}$ magyarázott el]]]]]
five-than more times explained
'The teacher explained several problems more than five times to every child'

It has always been known that different noun phrases occur in different positions, but their distribution had not been systematically investigated. Table (3) shows the results of my investigation for a rather representative sample of noun phrases. Some of the data are relevant only for a more detailed analysis, but I did not delete them because that might give a false impression of what noun phrases behave alike. Two points to be highlighted are as follows.

First, many noun phrases can occur in more than one preverbal position. I will argue that their interpretations vary accordingly.

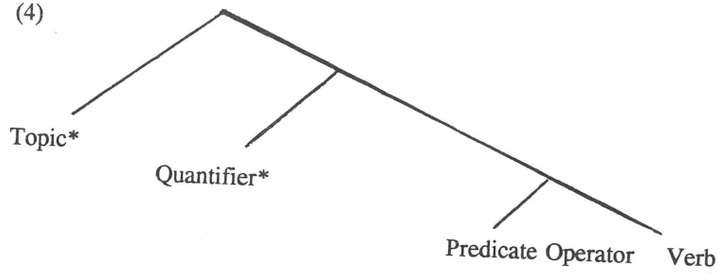
Second, let us single out a few noun phrases for each of the three positions that we are interested in. The Topic position accommodates names, definites, and unmodified indefinites. The Quantifier position accommodates universals, *is* 'even/also' phrases and *több, mint n* 'more than n' phrases (of one morphological variety). This last type also occurs in Predicate Operator; other predicate operators are noun phrases with determiners like *kevesebb, mint n* 'less than n' and *pontosan n* 'exactly n.' These data are illustrated in the simplified structure (4).

(3)

	Topic	Quant.	Focus	PredOp	Post-V
a legtöbb fiú 'most (of the) boys'	+				+
valamely fiú/bizonyos fiúk 'some boy(s)'	+				+
Péter, Péter és Mária 'Peter,' 'P and M'	+		+		+
a fiú(k) 'the boy(s)'	+		+		+
hat fiú 'six boys'	+		+	+@	+
sok fiú 'many boys'	+	+		+@	+
minden fiú 'every boy'		+			+
valamennyi fiú 'each boy'		+			+
még Péter is 'even Peter'		+			+
Péter is 'Peter, too'		+			+
semelyik fiú (neg.concord!) 'none of the boys'		+			+
legalább hat fiú 'at least six boys.'		+			+
több, mint hat fiú 'more than six boys(anal)'		+		+	+
hatnál több fiú 'more than six boys(syn)'				+	+#
pontosan hat fiú 'exactly six boys'				+	+#
kevés fiú 'few boys'				+	+#
kevesebb, mint hat fiú				+	+#
hatnál kevesebb fiú 'less than six boys(anal,syn)'				+	+#
legfeljebb hat fiú 'at most six boys'				+	+#
fiú(k) 'boy(s), existential'					

@ With noun destressed
# Only if Focus/PredOp is filled or V is negated

(4)



'John'	'every man'	'less than two men'
'the man'	'even/also two men'	'exactly two men'
'two men'	'more than two men'	'more than two men'

3. Distributive and collective readings

One sharp interpretive difference between these positions concerns the availability of distributive and collective readings. The use of the telic predicate 'lift up' is crucial in getting a clear picture of the data. The annotations Topic, Quant, and PredOp indicate the position of the subject noun phrase of the sentence:

(5) Topic Kati és Mari
Hat fiú fel-emelte az asztalt.
'Kati and Mari
Six boys lifted up the table'
collective ✓, distributive ✓

Quant Minden fiú
Több, mint hat fiú fel-emelte az asztalt.
Hat fiú is
'Every boy
More than six boys lifted up the table'
As many as SIX boys
*collective *, distributive ✓*

PrdOp Kevesebb, mint hat fiú
Több, mint hat fiú emelte fel az asztalt.
'Less than six boys
More than six boys lifted up the table'
collective ✓, distributive ✓

The first observation is that while both Topic and PredOp support both collective and distributive readings, Quantifier only supports distributive ones. This is quite understandable in the case of 'every' phrases but it is quite spectacular in the case of the others. 'More than n' phrases occur in both Quantifier and PredOp, and they differ accordingly with respect to their ability to support distributive readings. Also, in *hat fiú is* 'as many as SIX boys' the particle *is* 'even' adds nothing but the implicature that six is considered many; nevertheless, the noun phrase can now only occur in the Quantifier position and can only be interpreted distributively.

This suggests that the Quantifier position is one that is directly associated with a distributive operator. In fact, Stowell and Beghelli (1994, 1995) argue that universals appear in such a position at LF in English. The Hungarian data, which involves a much wider range of noun phrases, supports the assumption of such a distinguished position.

We may now ask if the collective and the distributive readings are obtained in the same way in Topic and in PredOp. The answer seems to be No. In the case of Topic phrases, the noun phrase denotes a potentially plural individual which may then be the subject of either collective or distributive predication. This scenario makes no intuitive or technical sense in the case of PredOp. Here we must start with the denotation of the predicate 'lifted up the table,' which may be distributive (contain sets of individuals) or collective (contain plural individuals). The contribution of the subject, then, is to say that the elements/atoms of the predicate denotation are boys, and to count those elements/atoms.

This means, for instance, that the PredOp variant of (5), 'More/fewer than six boys lifted the table' can have a collective reading, although the subject never denotes a collective (a plural individual). Note, by the way, that this sentence has the flavor 'It took more/fewer than six boys to lift the table (because it was so heavy/light),' but not all sentences of this form do. E.g. (6) has no such flavor.

(6) More/fewer than six estates surround this castle.

4. Topic versus PredOp: two modes of operation

Let us leave the Quantifier position for a while and focus on Topic and PredOp. You will have noticed that my paraphrases above already involve the "modes of operation" perspective mentioned in the overview. Namely, in the case of Topics, we start out with an entity associated with

the noun phrase and let the (relevant part of the) rest of the sentence be predicated over it. In the case of PredOps, we start out with the denotation of the predicate and let the noun phrase perform an operation (I specified counting) on it. The first is called the predicational and the second the counting or predicate operator mode. I believe these are two distinct canonical verification procedures.

At this point it is obvious that the noun phrases that operate in the predicational mode are those that DRT assumes introduce discourse referents. (There is one potentially divergent case, the noun phrase with determiner *a legtöbb* 'most,' which I do not wish to make specific claims about here.) On the other hand, the noun phrases that I call predicate operators behave, in Kamp & Reyle's terminology, as true generalized quantifiers.

I believe that the modes of operation perspective and the DRT perspective are not in conflict. Rather, I am proposing to add a procedural flavor to DRT. As Hans Kamp (p.c.) noted, DRT is already procedural in the sense that it matters what procedures create the DRSs, even when those DRSs are truth-conditionally equivalent. The verification procedures I have in mind go beyond this; they pertain to the interface between DRSs and the model theory.

In any case, I believe that slightly modified DRSs are suitable for serving as a representational basis for the procedural distinctions. Therefore, if the reader wishes to stick with classical DRT, (s)he may read the following as an argument for interpreting the distinguished positions of Hungarian in terms of DRT.

One slight modification that I am proposing pertains to the treatment of Topics. Kamp & Reyle stipulate that all (and only) noun phrase internal conditions are entered at the point where the discourse referent is introduced. This stipulation is in fact crucial in eliminating a problem noted by Heim (1982):

(7) If a cat likes a friend of mine, I always give it to her.

This sentence has a reading on which it is about my friend Sylvia. But if it were formalized by 'there is an entity such that if it is a friend of mine and a cat likes her, then I give the cat to her,' then it might be verified by any model that contains an entity that is not a friend of mine. This issue has received much attention recently in Abusch's and Reinhart's work. Abusch proposes a syntactic mechanism to percolate the restriction up to the existential quantifier, while Reinhart proposes to interpret indefinites like *a friend* as *f(friend)*, where *f* is a choice function. The results seem

to be equivalent to Kamp & Reyle's.

My own proposal is to stick with Kamp & Reyle's solution but to make it perhaps a little more principled. Instead of assuming, as is usual in all versions of DRT, that discourse referents are variables that range over the whole universe and the restrictor is predicated of them, I propose to semantically restrict, i.e. sort, the variables in an appropriate way. Specifically, I propose that discourse referents, i.e. logical subjects of predication, are variables ranging over witness sets of the GQ the noun phrase denotes. (Recall: a witness set is an element of the GQ that is also a subset of the smallest set GQ lives on. E.g., a witness of *two men* is any set that contains two men and no non-men.)

This proposal may be somewhat unusual in that it appeals directly to the denotation, as opposed to the syntactic form, of the noun phrase. But the Kamp and Reyle version of DRT is already more semantic than the pioneering versions were, e.g. in that it contains various operations (abstraction, summation) that create discourse referents but do not simply correspond to a syntactic operation in the sentence.

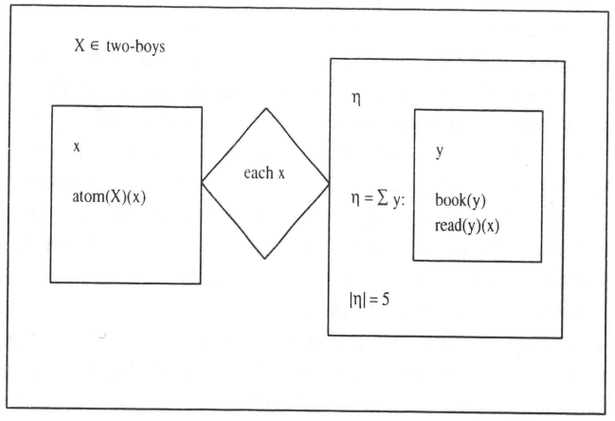
This modification makes a prediction: what is relevant for discourse referents is their witness sets and not their restrictor sets (i.e. not just the sets they live on). This seems to be supported by the following consideration. It has often been suggested that specific/topic indefinites are presuppositional in some sense. The authors who make this claim, e.g. Diesing (1992), typically mean that the restrictor set is presupposed to be non-empty. On my proposal, the source of presuppositionality may be the fact that predication is not merely false but nonsensical if there is no subject for it. But what is presupposed? Consider,

(8) Hat lengyel pápa könyvet írt.
six Polish Popes book-acc wrote
'Six Polish Popes wrote books'

On the usual approach, this sentence was infelicitous say 50 years ago, because the set of Polish Popes was empty, and it is felicitous now, because the set of Polish Popes is not empty. But it seems to me the mere existence of one Polish Pope does not make it felicitous. In order for it to be, we ought to have six Polish Popes. This is exactly what the witness set approach predicts.

So, the somewhat modified DRSs that I propose to adopt for Topics and PredOps are as follows.

(9) $[_{\text{Topic}}$ two boys $[_{\text{PredOp}}$ exactly five books $[_{\text{Verb}}$ read]]]



X∈DP is a variable ranging over plural individuals whose atoms are the elements of some minimal witness set of ⟦DP⟧. I follow Stowell and Beghelli in identifying the distributive operator with silent floated *each*. The formalization of the contribution of PredOp remains as in K&R (assuming an extension to cover the collective readings, discussed above). K&R's abstraction operator Σ forms the intersection of the sets of books and things read by x.

The assumption that the referent X introduced by a Topic is based on a minimal witness is motivated by collective readings and anaphora facts. A non-minimal witness of ⟦two boys⟧ may be a set of four boys. But if the table was lifted up by a collective of four boys, then 'Two boys lifted up the table' is false. (I thank Y. Winter for discussion on this point.) When the predicate is distributive, as in (9), there is no truth conditional difference between using minimal and non-minimal witnesses, thanks to the fact that all quantifiers in this position are monotonic increasing (a fact to be commented on below). For uniformity, then, we want to assume that the referent is based on a minimal witness regardless of what kind of predicate it will subsequently be a subject of, and anaphora facts confirm that this is indeed correct:

(10) Two boys came in. They were tired.

While the first sentence is compatible with, say, four boys coming in, they only refers back to the two that were singled out.

5. Where does the Quantifier position belong?

Let us now move on to the question of how noun phrases that occur in the Quantifier position are to be analyzed. Specifically, where do they stand in the dichotomy we are considering? DRT classes them with generalized quantifiers; I will suggest that they in fact side with the noun phrases that operate in the predicational mode and introduce discourse referents.

In making this argument, the Hungarian data are crucial, for two reasons. One is intuitive, the other is technical.

As regards the intuitive aspect, consider the fact that there are noun phrases that occur in both the Quantifier and the PredOp positions. These are 'many' and 'more than n' phrases. E.g.,

(11) a. Sok diákunk megbukott. Quant
'Many students of ours flunked'
b. Több, mint hat diákunk megbukott.
'More than six students of ours flunked'

(12) a. Sok diákunk bukott meg. PredOp
Több, mint hat diákunk bukott meg.

I chose the examples in such a way that, due to the possessive construction, they are both "presuppositional" and the 'many' phrases are interpretable as proportional in both cases. If this is so, then there is no standardly known reason for the sentences in (11) and (12) to be perceived as not meaning the same. But that is the perception; no native speaker would be tempted to say otherwise, even though they might not be able to explicate the difference. So this is something to account for.

My account is that in (11) we take a set of students and claim that each of them flunked. In (12), we take those who flunked, and count our students among them. (11), but not (12), is "about" a set of students. A further difference, as László Kálmán observes is that (12) presupposes that the issue of people flunking is under discussion. This meshes with the proposed account and is related to the fact that PredOp is a kind of "weak focus," cf. diagram (1).

Now we can go on to check whether noun phrases in Quantifier exhibit the property that I claimed was the DRT counterpart of acting as a logical subject of predication, that is, whether they introduce discourse referents.

On this technical side, the Hungarian data are useful, because they

make various arguments possible that universals do not. 'Every man' type phrases are principal filters. This means they have a unique witness set that is identical to their restrictor set. This circumstance makes it impossible to check whether they can exhibit non-maximal reference anaphora, which is one of the marks of introducing a discourse referent in DRT. It also makes it impossible to check whether they exhibit referential dependence. The 'many' and 'more than n' phrases in Hungarian considered above make these tests possible.

At the outset, it is to be noted that noun phrases in the Quantifier position clearly do not introduce the same type of discourse referent as, say, definites. Kamp & Reyle's basic test is as follows:

(13) Two lawyers hired a secretary who they had interviewed.

(14) More than five lawyers hired a secretary who they had interviewed.

In (13), it is possible to interpret *hired a secretary who they had interviewed* distributively (i.e. each lawyer hires a different secretary) and the pronoun *they* inside it, collectively. This shows unambiguously that *two lawyers* introduces a plural individual referent that antecedes *they* on this interpretation. But no such interpretation is possible with *every lawyer* or *more than five lawyers*. The same contrasts obtain in Hungarian, between Topic and Quantifier.

This in itself only shows that if the noun phrases in Quantifier position introduce discourse referents, they cannot be plural individuals. But there are in fact at least two pieces of evidence that they introduce a referent of some sort.

One kind of evidence comes from the following contexts:

(15) Quant Több, mint hat diákunk megbukott.
Rosszul tudod, átengedték őket.
'More than six students of ours flunked.
You are not well-informed, they passed'

PrdOp Több, mint hat diákunk bukott meg.
* Rosszul tudod, átengedték őket.
'More than six students of ours flunked.
* You are not well-informed, they passed'

When 'more than six students of ours' is in PredOp, the counting position,

the dialog is incoherent. The second sentence can only mean that all the students who flunked passed. This is in fact expected if noun phrases in PredOp do not introduce a referent and thus only support maximal reference anaphora (that is, anaphora to the intersection of the sets of our students and those who flunked). But when the noun phrase is in Quantifier position, all is well, there is no incoherence. At first blush one might think the reason is that the second sentence is understood as, 'All the students who you imagine flunked passed.' But if this were the case, this would be a way-out in the PredOp dialog, too. The only other alternative is to assume that in Quantifier position, 'more than six students of ours' introduces a set, and 'they' refers back to that set. Now there is no room for contradiction, since the definition of this set did not involve the flunking property.

The second, perhaps stronger piece of evidence is this. Consider:

(16) Quant Legalább két elemző úgy gondolja, hogy <u>több, mint hat hazug</u> igazat beszél.
✓ 'there are <u>>6 liars</u> who a fixed set of ≥ 2 analysts thinks speak the truth'

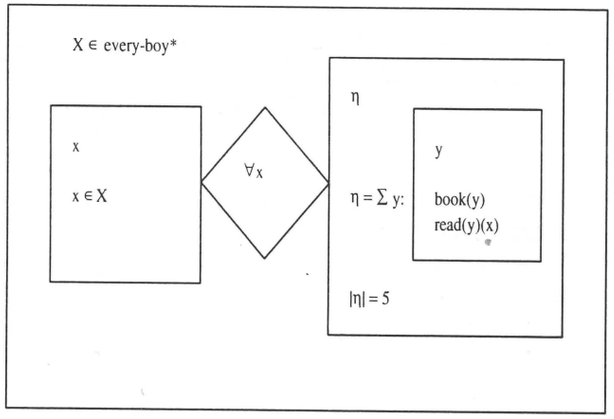
PrdOp Legalább két elemző úgy gondolja, hogy <u>több, mint hat hazug</u> beszél igazat.
* 'there are <u>>6 liars</u> who a fixed set of ≥ 2 analysts thinks speak the truth'
✓ 'each of ≥ 2 analysts thinks that the # of liars who speak the truth is >6'

In both cases, the liars can be liars either in the speaker's world or in the analysts' worlds. But there is a contrast concerning a further matter. When 'more than six liars' is in PredOp, the sentence can only mean that at least two analysts think that the number of liars who speak the truth is greater than six. The sentence can in no way be interpreted as being about a fixed set of liars that all the analysts have some belief about. On the other hand, when 'more than six liars' is in Quantifier position, it does have such an interpretation. This latter is exactly what we predict based on Kamp & Reyle's treatment of discourse referents. First, a discourse referent can be introduced into any superordinate box; this is why 'more than six liars' need not be dependent on 'at least two analysts.' But when this happens, the distributive operator associated with the referent invariably gets stuck in its "base" position. This is why the analysts also form an independent set.

These two pieces of data suggest that a noun phrase in Quantifier position indeed introduces a referent and there is a separate distributive operator associated with it. I will assume that this referent differs from more standard referents in that it is a set, not a plural individual referent, and whatever differences exist in their behavior follows from this.

Recall that universals occur in the same Quantifier position. Since their behavior is fully compatible with the above reasoning, I will assume that they, too, introduce the same kind of set referents. Thus the modified DRT-style representation is as below. Notation: $X \in DP^*$ is a variable ranging over witness sets of $[\![DP]\!]$.

(17) $[_{Quant}$ every boy $[_{PredOp}$ exactly five books $[_{Verb}$ read]]]



Note that here the referent is not required to be a minimal witness. Since these sentences only have distributive readings, and since the Quantifier position hosts only increasing quantifiers (see below), truth conditions will not motivate the decision. But anaphora facts do. Recall (15), 'More than six students of ours flunked. You are not well-informed, they passed.' If the referent I argue 'more than six students of ours' introduces were a minimal witness set of the quantifier, this discourse would make sense only if 'they' referred back to exactly seven students. But this is not necessary; eight or more students may well be involved.

Let me now cite two pieces of circumstantial evidence in favor of the claim that inhabitants of the Quantifier position introduce a set referent with a distributive operator attached.

One piece of evidence comes from the fact that Stowell and

Beghelli's analysis of the Logical Form behavior of the English universal *every man* amounts to the same claims. Recall that the Hungarian Quantifier position is claimed to be analogous to the position of English universals, only it is better populated. Specifically, two of the phenomena S&B are concerned with are as follows:

(18) John didn't read every book.

(19) What did every boy read?
John wonders what every boy read.

The notable property of (18) is that, on normal intonation, it only allows a reading where *not* takes scope over *every book*. The notable property of the sentences in (19) is that they have pair-list readings. S&B analyze both cases by assuming that the universal acts as a variable bound by some operator (the negation or the question operator). Details aside, this would make no sense on the usual interpretation of universals, but it makes good sense if the universal introduces a set referent, since that is a bindable variable in DRT terms. (Incidentally, the result that universals may be bound is not unique for this analysis; dynamic semantics can produce the same, as observed by Groenendijk & Stokhof (1993).)

The other piece of circumstantial evidence pertains to an analysis of pair-list readings that this audience will be more familiar with. Recall that Groenendijk and Stokhof, Higginbotham, and Chierchia all analyzed pair-list readings in such a way that the universal does not act in its usual manner; rather, its unique witness set serves to restrict the domain of the question. A highly simplified form of Chierchia's analysis is below:

(20) What did every boy read?
$\lambda P \exists W[witness(W, [\![\text{every boy}]\!]) \ \& \ P(\text{which } w \in W \text{ read what})]$

That is, Chierchia's analysis factors the contribution of *every boy* into a witness set and a distributive operator. This is precisely what Stowell and Beghelli do and what I argued for above. The only difference is that Chierchia assumes that *every boy* exhibits this behavior only under exceptional circumstances, i.e. in pair-list readings, while I argue that this is its standard behavior.

6. Denotational semantics and modes of operation

This brings us to the last issue that I wish to raise in this paper, namely, to what extent do the denotational semantic properties of a noun phrase determine its mode or modes of operation.

Some noun phrases have no chance to operate in the predicational mode (or, to introduce discourse referents). Assuming that Det is conservative and has extension, observe:

(21) If Det is increasing, but not if it is decreasing or non-monotonic, Det(N)(P) = for some witness W of Det(N), for every $x \in W$, P(x)

The left hand side is the traditional interpretation, and the right hand side is the one that involves grabbing a witness and predicating over each of its elements. In view of (21), the latter is a possible way to construe to operation of increasing quantifiers, but if decreasing or non-monotonic quantifiers were construed in the same way, that would seriously distort their truth conditions. (Note: *Fewer than five men walk* and *Exactly five men walk* do not just mean that there is a set with the said cardinality whose members are men and walk; they insist that this is a maximal set.)[2]

So, the question is, do the Topic and Quantifier positions conform to this requirement? Going back to table (3) we can see that they do!

(22) All the decreasing and non-monotonic quantifiers of Hungarian are confined to the PredOp position.

The restriction that Topic and Quantifier accommodate only increasing quantifiers is a robust fact that certainly calls for an explanation. I am not aware of any possible explanation other than the one that is emerging here, namely, that the inhabitants of these positions operate in a fashion that is available only to increasing quantifiers.

We must note, however, that PredOp also contains some increasing quantifiers. What this indicates is that the abstract denotational semantic ability of a noun phrase to introduce a discourse referent does not force it to do so. That is, denotational semantics delimits but does not determine actual modes of operation.

---

[2] Exactly the same reasoning applies to the standard analyses of pair-list readings, cf. (20). In Szabolcsi (1994) I observed that in fact, in certain complement contexts non-increasing quantifiers also suppprt pair-list readings, wherefore that analysis needs to be rejected *qua* generally valid solution. In the previous section I was using (20) as supporting evidence strictly in connection with universals.

This conclusion is reinforced by subtle data like the following. The 'more than' construction has two versions in Hungarian, morphological and syntactic comparison. The two seem denotationally equivalent. But while the one involving syntactic comparison can occur in either Quantifier or PredOp, the one involving morphological comparison can only occur in PredOp.

A quite similar conclusion may be drawn for the other mode of operation. Recall that I claimed that the noun phrases that act like true generalized quantifiers are actually all "counters." This is not to say that they are necessarily intersective, nevertheless, they all involve a counting (and perhaps a comparing) procedure that gives a number in the end.

Universals and definites clearly do not conform to this pattern and they do not occur in PredOp, indeed. But what happens in the converse situation: is the abstract ability of a quantifier to count suffice for it to operate in the counting mode?

I have not analyzed the behavior of the determiner *a legtöbb* 'most' in detail, but looking back at table (3) we may note that it is quite sharply restricted to occurring in the Topic position. But consider the fact that *Most of the boys left* is considered to be synonymous with *More than 50% of the boys left*. Now, it turns out that the Hungarian noun phrase 'more than 50% of the boys' can occur in PredOp, in distinction to 'most of the boys.' So here, too, we see that putative denotational synonymy does not predestine two noun phrases to operate identically.

Interestingly, Hungarian word order is not the only empirical domain that sets these two noun phrases apart. Consider *there*-insertion and binominal *each* contexts in English:

(23) a. There will be more than 50% of the boys in the yard.
b.* There will be most of the boys in the yard.

(24) a. The boys read more than 50% of the books each.
b.* The boys read most of the books each.

Sutton (1993), whose work is the source of these data concluded, somewhat desperately, that the noun phrases that combine with binominal *each* do not seem possible to characterize in denotational terms; rather, what they seem to have in common is that they are all "counters." While the proposal I made above does not immediately explain why specifically counters need to be involved in this construction, I hope to have substantiated that this type of non-denotational semantic conclusion need not be that desperate.

References

Abusch, D.: 1994, The scope of indefinites, *Natural Language Semantics* 2.

Beghelli, F.: 1995, The syntax of distributivity and pair-list readings, in Szabolcsi, A. (ed.), *Ways of Scope Taking*. Kluwer (forthcoming).

Beghelli, F. and Stowell, T.: 1994, The direction of quantifier movement. GLOW lecture, Vienna.

Beghelli, F. and Stowell, T.: 1995, Distributivity and negation, in Szabolcsi, A. (ed.), *Ways of Scope Taking*. Kluwer (forthcoming).

Chierchia, G.: 1993, Questions with quantifiers, *Natural Language Semantics* 1.

Diesing, M.: 1992, *Indefinites*. MIT Press.

Heim, I.: 1982, *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst.

Groenendijk, J. and Stokhof, M: 1993, Interrogatives and adverbs of quantification, in Bimbó, K. and Máté, A. (eds.), *4th Symposium on Logic and Language*. Akadémiai.

Kamp, H. and Reyle, U.: 1993, *From Discourse to Logic*. Kluwer.

É. Kiss, K.: 1991, Logical structure in linguistic structure: the case of Hungarian, in Huang, J. and May, R. (eds.), *Logical Structure and Linguistic Structure*. Kluwer.

Reinhart, T.: 1995, *Interface Strategies*. Unpublished manuscript, Utrecht University.

Sutton, M.: 1993, *Binominal each*. MA thesis, UCLA.

Szabolcsi, A.: 1994, Quantifiers in pair-list readings and the non-uniformity of quantification, in Dekker, P. and Stokhof, M. (eds.), *Proceedings of the 9th Amsterdam Colloquium*. ILLC, Amsterdam.

Szabolcsi, A. (ed.): 1996, *Ways of Scope Taking*, Kluwer (forthcoming).

A Critique of a Proof-Theoretic Account of Anaphora

Matthew Watson, University of Texas at Austin

Recently, Aarne Ranta has presented a proof-theoretic framework, based upon Martin-Löf's intuitionistic type theory, for studying anaphora in English discourse (Ranta 1991 and Ranta 1994). Ranta has claimed that this proof-theoretic perspective is superior in some respects to model-theoretic treatments of discourse anaphora, such as Discourse Representation Theory (hereafter DRT) and Dynamic Predicate Logic (hereafter DPL). His main argument for this claim centers around the analysis of sentences such as the following:

(1) If every man finds a pen some man loses it.

Both DRT and DPL say that the pronoun 'it' cannot be anaphorically linked to the indefinite noun phrase 'a pen'; Ranta's theory, as we shall see, says that the pronoun 'it' refers to the pen found by the man introduced by the noun phrase 'some man'. Ranta defends this result as a proper analysis of sentences such as (1). I will attempt to cast some doubt on his contention. I will first argue that the anaphoric links which Ranta alleges to be present in sentences such as (1) are not clearly present there. I will then argue that even if Ranta is correct about the presence of these links, his analysis of the meaning of these sentences is not clearly the correct one. After presenting these considerations against Ranta's major argument for the superiority of his proof-theoretic approach over the more established model-theoretic approaches, I will spend the remainder of the paper pointing to similarities between the respective approaches. To be more specific, I will argue that there is a similarity between the classical DRT reading of donkey sentences with determiners and relative clauses and Ranta's reading of these sentences. Because the readings given to these sentences by the respective approaches are so similar, we can find analogs of the problems associated with the classical DRT reading (most notably the proportion problem) in Ranta's reading. I will conclude this paper by suggesting how these problems might be repaired.

Before looking at any specific analyses of sentences within Ranta's framework, I will briefly present its general structure. The formal framework which Ranta employs to study the semantics of English anaphoric constructions is a kind of categorial grammar, consisting of three parts. One of these parts is a lexicon, which consists of "category assignments to *basic expressions* " (Ranta 1991, p.227). In the lexicon certain English words are categorized as expressions standing for sets, others are categorized as expressions standing for propositional functions over those sets, and so on. For instance, the common noun 'farmer' is given the following categorization in the lexicon:

(2) farmer: set.

'farmer: set' is the judgment that the type of farmers is a set. This is a judgment in Ranta's extension of Martin-Löf's higher level type theory (for a full account of this theory, see Ranta 1991 or Nordström, Petersson, and Smith 1990). Every entry in the lexicon is a judgment in this theory; the form of judgment used for a particular lexical entry depends upon the grammatical category to which the lexical item belongs. Another component of Ranta's formal framework is what he calls the "*categorial grammar* " (Ranta 1991, p.227), which simply consists of the rules of Martin-Löf's higher level type theory. These rules are used to build up complex propositional expressions from the basic expressions of the lexicon. The final component of the framework is a set of "*sugaring rules* " (Ranta 1991, p.227), which transform propositional expressions of the type theory into strings of English words (sugaring rules can be thought of as the inverse of parsing rules). These rules are not one-to-one; that is, there are some English expressions which can be obtained by applying the sugaring rules to two or more distinct propositions of type theory. Moreover, the rules are not even functional, since the same proposition may be "sugared" into more than one English expression. The rules are meant to be meaning-preserving, so that an English expression obtained by sugaring a propositional expression of the type theory has the same meaning as that expression (consequently, a single English expression which can be obtained by sugaring from distinct propositions of type theory is ambiguous). The meanings of propositions in the higher level type theory (and, consequently, the meanings of the English expressions into which they can be sugared) are given by Martin-Löf's meaning explanations for judgments in that theory.

The following judgment can be derived in Ranta's version of higher level type theory:

(3) (Πx: (Πy: man)(Σz: pen)find(y,z)) (Σu: man)lose(u,p(ap(x,u))): prop.

The proposition (Πx: (Πy: man)(Σz: pen)find(y,z)) (Σu: man) lose(u, p(ap(x,u))) can then be sugared into the English sentence (1). The derivation of this proposition in the higher level type theory gives some insight into its structure:

1.	2.	3.	4.
x: man	y: pen find	z: (Πx: man)(Σy: pen)find(x,y)	u: man ap
	find(x,y): prop abstr. (2)	4.	ap(z,u): (Σy: pen)find(u,y) p
pen: set	(y)find(x,y): (pen)prop Σ	u: man	p(ap(z,u)): pen lose
	(Σy: pen)find(x,y): prop abstr. (1)		lose(u,p(ap(z,u))): prop abstr. (4)
man: set	(x)(Σy: pen)find(x,y): (man)prop Π	man: set	(u)lose(u,p(ap(z,u))): (man)prop Σ
	(Πx: man)(Σy: pen)find(x,y): prop		(Σu: man)lose(u,p(ap(z,u))): prop abstr.(3)
			(z)(Σu: man)lose(u,p(ap(z,u))): ((Πx: man) (Σy: pen)find(x,y))prop Π

(Πz: (Πx: man)(Σy: pen)find(x,y)) (Σu: man)lose(u,p(ap(z,u))): prop.

The left side of this derivation is a derivation of the judgment that (Πx: man)(Σy: pen)find(x,y) is an object of the type of propositions. This derivation uses certain judgments given in the lexicon (man: set, pen: set, find: (man)prop, and so on), as well as a number of rules in Martin-Löf's higher level type theory. These rules are written to the right of the lines of the derivation to which they are applied. Note that some of these rules (find, lose, Σ, Π, ap, and so on) are not basic rules of higher level type theory, but can be obtained by applying the basic rules to the categorizations of the corresponding constants (find, lose, Σ,Π, ap, and so on) given in the lexicon. Also note that whenever the abstraction rule (abstr.) is used to form an object of type (A)B (the type of functions from A to B), all or some of the occurrences of the hypothesis x: A are discharged; this is indicated in the derivation by writing the numeral corresponding to the hypothesis discharged in an application of this rule to the right of the line to which the rule is applied.

The right side of the derivation given above begins with the hypothesis that z is an object of the type (Πx: man)(Σy: pen)find(x,y) and the hypothesis that u is an object of the type man. An object of type (Πx: A) B(x) is a function which, when applied to a member a of the set A, yields a member of the set B(a). So when z is applied to u, this yields the application instance ap(z,u): (Σy: pen)find(u,y) (actually, the constant ap defined in the higher level type theory also takes the types A and B as arguments; following Ranta 1991, I am suppressing the type information). The type (Σy: pen)find(u,y) is a set whose canonical members are pairs (a,b), where a: pen and b: find(u,a). Since ap(z,u) is a member of this set, it can be computed into canonical form (that is, it can be computed into a pair satisfying the condition mentioned above). Consequently, the projections p(ap(z,u)): pen and q(ap(z,u)): find(u,p(ap(z,u))) can be formed. In the derivation above, the first of these projections is derived from ap(z,u): (Σy: pen)find(u,y). Since this projection is a member of the set pen, and since lose is a propositional function defined on the sets man and pen, the proposition lose(u,p(ap(z,u))) can be formed. The variables in this expression can then be bound by applying the Σ and Π rules in the order indicated above. Both the constant Σ and the constant Π take a set and a propositional function defined on that set as arguments and yield a proposition as a value. Since sets and propositions are identified in Martin-Löf's type theory (through a version of the Curry-Howard isomorphism), the final application of the Π rule in the derivation is correct.

When the proposition (Πz: (Πx: man)(Σy: pen)find(x,y))(Σu: man) lose(u,p(ap(z,u))) is sugared into the sentence (1), the term p(ap(z,u)) is sugared into the pronoun 'it'. As I noted above, p(ap(z,u)) is a member of the set pen. Moreover, this pen is the pen found by the man introduced by

the indefinite noun phrase 'some man', since q(ap(z,u)): find(u,p(ap(z,u))) (that is, q(ap(z,u)) is a proof object of the proposition find(u,p(ap(z,u)))). So even without giving the full details of the type-theoretic account of the meaning of the proposition (Πz: (Πx: man)(Σy: pen)find(x,y)) (Σu: man) lose(u,p(ap(z,u))), and hence of the meaning of sentence (1), it is clear that the pronoun 'it' and the indefinite noun phrase 'a pen' are predicted by Ranta's theory to be anaphorically linked.

Both DRT and DPL predict that 'it' and 'a pen' cannot be anaphorically linked in sentence (1). Ranta views this as a limitation of those theories, since he thinks that there is such a link in this sentence. However, it is not clear that there is such a link. If there is such a link, then one would expect that a similar link would be present in other sentences with the same form. But it is not clear, for example, that the pronoun 'it' can take the noun phrase 'a donkey' as its antecedent in the following sentence:

(4) If every farmer beats a donkey some farmer harms it.

Both DRT and DPL predict that 'a donkey' cannot serve as an antecedent for 'it' in this sentence, for the same reasons that they predict that 'a pen' cannot serve as an antecedent for 'it' in sentence (1). But Ranta's theory says that 'a donkey' and 'it' can be anaphorically linked in sentence (4), since (4) can be sugared from the following proposition of type theory:

(5) (Πz: (Πx: farmer)(Σy: donkey)beat(x,y)) (Σu: farmer) harm(u,ap(z,u)).

Here p(ap(z,u)): donkey and q(ap(z,u)): beat(u,p(ap(z,u))). It is even less clear in the following sentence than in sentence (4) that the noun phrase 'a donkey' can serve as an antecedent for the pronoun 'it':

(6) If every veterinarian knows a farmer who owns a donkey some veterinarian vaccinates it.

Again, DRT and DPL predict that 'a donkey' and 'it' cannot be anaphorically linked in (6). But (6) can be sugared from the following proposition of type theory:

(7) (Πv: (Πu: veterinarian)(Σz: (Σx: farmer)(Σy: donkey)own(x,y)) know(u,p(z))) (Σw: veterinarian)vaccinate(w,p(q(p(ap(v,w)))))).

Here p(q(p(ap(v,w)))): donkey, since p(ap(v,w)): (Σx: farmer)(Σy:donkey) own(x,y), and hence q(p(ap(v,w))): (Σy: donkey)own(p(p(ap(v,w))),y). Since q(q(p(ap(v,w)))): own(p(p(ap(v,w))), p(q(p(ap(v,w))))), Ranta's analysis of (6) predicts a link between 'a donkey' and 'it' in that sentence. If I am correct concerning the dubiousness of the anaphoric links predicted by Ranta's theory in examples such as (4) and (6), then Ranta's main argument for the superiority of his theory as a tool for analyzing discourse anaphora is undermined. At the very least, Ranta would have to provide some principled way of distinguishing between such cases as (1), where the predicted anaphoric link is arguably acceptable, and such cases as (6), where the predicted link is arguably unacceptable. Perhaps this could be

done by introducing some constraints on such links into the existing theory. However, any discussion of the nature of these constraints, or even where such constraints should be placed within Ranta's three-part theory (e.g., whether they should be placed in the categorial grammar or the sugaring rules) is beyond the scope of this paper.

Even if one accepts as genuine the anaphoric links which Ranta predicts to hold in sentences (1), (4), and (6), it is possible that his theory provides an incorrect analysis of the meanings of those sentences. I will explore this possibility by looking in detail at the meaning assigned by Ranta's theory to sentence (4). The meaning assigned by the theory to this sentence is identical to the type-theoretical meaning of the proposition from which it is sugared (proposition (5)). The meaning of a proposition in Martin-Löf's type theory is given by explaining what constitutes a canonical proof of that proposition, as well as explaining under what conditions two canonical proofs of that proposition are equivalent. A canonical proof of proposition (5) is a function which, when applied to a proof a of the proposition (Πx: farmer)(Σy: donkey)beat(x,y), yields a proof object b(a) of the proposition (Σu: farmer)harm(u,p(ap(a,u))) (where b(z): (Σu: farmer)harm(u,p(ap(z,u))), given the hypothesis that z: (Πx:farmer)(Σy: donkey)beat(x,y)). Since b(a) is a proof object of a proposition of the form (Σx: A)B, it can be computed into a pair (c,d). The first member of this pair is an element of the set of farmers,while the second is a proof of the proposition beat(c,p(ap(a,c))). The expression ap(a,c), when computed, yields a proof of the proposition (Σy:donkey) beat(c,y). When the projection p is applied to this proof (yielding p(ap(a,c))), a member of the set of donkeys is obtained. Putting this all together, a canonical proof of proposition (5) is a general method which transforms an arbitrary proof that every farmer beats a donkey into a proof that some farmer harms the donkey he beats. This provides a specification of the proof conditions (or, if we replace 'proof' with 'verification' in the preceding sentence, the verification conditions) of sentence (4) within the framework of Ranta's theory.

In order to see whether or not the proof (or verification) conditions assigned by Ranta's theory to sentence (4) are correct, consider the following situation. Suppose there are only two farmers, John and James, and four donkeys, Beavis, Betty, Josie, and Magilla. Further suppose that there is a general method f which yields proofs that John beats Beavis and James beats Magilla, and a general method g which yields proofs that John beats Josie and James beats Betty. It follows that both f and g are proof objects of type (Πx: farmer)(Σy: donkey)beat(x,y). Finally, suppose that there are proofs that John harms Beavis and James harms Magilla, as well as proofs that any proof that John harms Josie, or any proof that James harms Betty, leads to absurdity. Ranta's theory says that sentence (4) cannot be proved (or verified) in this situation. For suppose that it could

be. Then there would be a general method f' which, when applied to an arbitrary proof a of the proposition (Πx: farmer)(Σy:donkey) beat(x,y), yields a proof b(a) of the proposition (Σu: farmer)harm(u, p(ap(a,u))). In particular, if f' were applied to g, this would yield a proof b(g) of the proposition (Σu: farmer)harm(u,p(ap(g,u))). But such a proof would have to be either a proof that John harms Josie or a proof that James harms Betty. By hypothesis each of these alternatives leads to absurdity. Consequently, (4) cannot, according to Ranta, be proved (or verified) in this situation.

The consequence of Ranta's theory presented in the preceding paragraph is not clearly correct. It is arguable that (4) has a proof (or verification) in the hypothesized situation. This is because there is a general method which transforms any proof that every farmer beats a donkey into a proof that at least one farmer harms at least one of the donkeys he beats. This method consists simply of looking at the members of the set of proofs that every farmer beats a donkey (in the hypothesized situation this set is a surveyable totality, since it contains but two members, f and g) and then seeing if there is a proof of the proposition (Σu: farmer)harm(u,p(ap(a,u))) for some member a of this set. If any such proof is found, simply match up every proof that every farmer beats a donkey with each of the proofs found. If no such proof is found, then there cannot be a proof (or verification) of sentence (4). In the hypothesized situation, there are two distinct functions which transform any proof that every farmer beats a donkey into a proof that at least one farmer harms at least one of the donkeys he beats. One of these functions transforms any proof that every farmer beats a donkey into a proof that John harms Beavis, while the other transforms any proof that every farmer beats a donkey into a proof that James harms Magilla.

I have argued that there is a proof of sentence (4) in the situation I have presented. The reason that Ranta's theory implies that there cannot be a proof of (4) in this situation is that it requires that every proof of the antecedent of (4) be transformed into a proof that some farmer harms the very donkey which the proof of the antecedent tells us he beats. I have proposed that there is a proof of (4) when every proof of the antecedent of (4) can be transformed into a proof that some farmer harms a donkey which some proof of the antecedent of (4) tells us he beats. This proposal can obviously be generalized to any sentence of the same form as (4), including such sentences as (1) and (6). Since alternative readings of the proof conditions of sentences such as (4) exist (I have just given one), Ranta needs to argue that the reading of the proof conditions of these sentences which he gives is a plausible one for at least some contexts in which they are used. He should not be satisfied with the proof conditions which his theory assigns to sentences such as (4) simply because those conditions predict an anaphoric link which other theories of discourse anaphora (such as DRT and DPL) miss; he must further argue (as he has

not in Ranta 1991 or Ranta 1994) that those proof conditions are the correct ones for those sentences in a wide range of contexts.

So far, I have tried to blunt the force of Ranta's main argument for the superiority of his grammar by arguing that it is not clear that the anaphoric links which his theory predicts to hold in sentences such as (1), (4), and (6) are acceptable, and by arguing that even if those links are acceptable, it is not clear that Ranta's grammar captures the meaning of those sentences. I now wish to point out certain similarities between the interpretation of donkey sentences provided by Ranta's theory, and the interpretation of those sentences provided by the model-theoretic approaches to discourse anaphora (DRT and DPL) which he criticizes in his main argument. To be more specific, I wish to point out a similarity between the interpretation given by classical DRT to donkey sentences with determiners and relative clauses - sentences such as the following:

(9) Every farmer who owns a donkey beats it.

(10) Some farmer who owns a donkey beats it.

(11) Most farmers who own a donkey beat it.

- and the interpretation given by Ranta's grammar to such sentences. I choose to compare Ranta's grammar with DRT rather than DPL simply because I am more familiar with the former theory. But given the similarity between the readings of donkey sentences given in DRT and in 'classical' versions of DPL, much of what I say here should also be applicable to that theory.

Sentence (9) can be obtained from sugaring the following proposition of type theory:

(12) (Πz: (Σx: farmer)(Σy: donkey)own(x,y))beat(p(z),p(q(z))).

Here p(z): farmer and p(q(z)): donkey, where q(z): (Σy: donkey) own(p(z),y). The meaning assigned to sentence (9) by Ranta's grammar is identical to the meaning given to proposition (12) in Martin-Löf's type theory. A canonical proof of proposition (12) is a function which, when applied to an arbitrary proof a of the proposition (Σx: farmer)(Σy: donkey) own(x,y), yields a proof b(a) of the proposition beat(p(a),p(q(a))) (where b(z): beat(p(z),p(q(z))) can be formed in the context z: (Σx: farmer) (Σy: donkey)own(x,y)). A proof a of the proposition (Σx: farmer) (Σy: donkey)own(x,y) can be computed into a pair (c,d), where c is a farmer and d is a proof that he owns a donkey. The second member of this pair can be computed into a pair (f,g), where f is a donkey and g is a proof that c owns f. The objects c and f can be recovered from the proof a by applying p to a (to obtain c) and applying p to q(a) (to obtain f). Consequently, a canonical proof of (12) is a function which, when applied to an arbitrary proof that a pair consisting of a man and a donkey stands in the owning relation, yields a proof that the same pair stands in the beating relation. In short, Ranta's theory says that the determiner 'every' in

sentence (9) is to be interpreted as quantifying over proofs that farmer-donkey pairs stand in the owning relation.

The interpretation which Ranta's theory assigns to the determiner 'every' in sentence (9) is similar to the interpretation which classical DRT assigns to this determiner in this sentence (see Kamp 1981). Classical DRT says that (9) is true if and only if every farmer-donkey pair which stands in the owning relation also stands in the beating relation. So 'every' is interpreted as quantifying over farmer-donkey pairs which stand in the owning relation. This is close to the reading given by Ranta's theory, which interprets 'every' as quantifying over *proofs that* farmer-donkey pairs stand in the owning relation. The correspondence between the two approaches is even closer if it is stipulated that there can be only one canonical proof that a certain farmer a owns a certain donkey b. If this stipulation is made, then there can only be one proof that farmer a owns donkey b, since any proof that farmer a owns donkey b can be computed into a canonical proof that farmer a owns donkey b, and since any two proofs which reduce to the same canonical proof are identified in Martin-Löf's type theory. Since there can be only one proof that a farmer a owns a donkey b, the number of elements in the set (Σx: farmer)(Σy: donkey)
own(x,y) will be equal to the number of farmer-donkey pairs which can be proven to stand in the owning relation. So if this stipulation is made, then Ranta's theory interprets the determiner 'every' as quantifying over farmer-donkey pairs which can be proven to stand in the owning relation. This interpretation is indeed close to the classical DRT interpretation.

DRT says that sentence (9) is true if and only if every farmer-donkey pair which stands in the owning relation also stands in the beating relation. So DRT says that this sentence is true just in case every farmer who owns a donkey beats every donkey he owns. This reading of sentence (9) has been called the "strong reading" (in Chierchia 1990) or the "∀-reading" (in Chierchia 1992). It is contrasted with the "weak reading" (Chierchia 1990) or the "∃–reading" (Chierchia 1992) of sentence (9):

(13) Every farmer who owns a donkey beats a donkey he owns.

The strong reading of sentence (9) is the correct default interpretation of that sentence, although the weak reading may be better if more is known about the circumstances in which it is uttered (for an example in which the circumstances of utterance favor the weak reading, see Chierchia 1992, p.116). Ranta's grammar delivers a version of the strong reading of sentence (9); that is, Ranta's grammar says that there is a proof of (9) just in case there is a proof that every farmer who owns a donkey beats every donkey he owns. This version of the strong reading of (9) gives the correct default proof (or verification) conditions for it.

There are some donkey sentences with the determiner 'every' which favor the weak reading as the default reading. Such sentences include the following:

(14) Every person who has a credit card will pay his bill with it. (Robin Cooper, cited in Chierchia 1992)

(15) Every person who has a dime will put it in the meter. (Pelletier and Schubert 1989).

The strong reading of (15) says that the sentence is true if and only if every person who has a dime will put every dime he has into the meter. But this sentence is clearly true even if some people put some, but not all, of their dimes in the meter. Classical DRT assigns the strong reading to this sentence, and thus fails to provide the correct truth conditions for it. If I am correct in seeing a parallel between the readings of donkey sentences given by classical DRT and those given by Ranta's grammar, then Ranta's grammar should fail to provide the correct proof conditions for this sentence. This sentence can be sugared from the following proposition of type theory:

(16) (Πz: (Σx: person)(Σy: dime)has(x,y))put-in-the-meter(p(z), p(q(z))).

For the sake of simplicity, I have not bothered to provide a further type-theoretical analysis of the prepositional phrase 'in the meter', nor have I taken into consideration the tense of sentence (15) (for a treatment of tense within type theory, see Ranta 1994, Chapter 5). Omitting these factors should not affect my argument. A canonical proof of proposition (16) is a function which, when applied to an arbitrary proof a of the proposition (Σx: person)(Σy: dime)has(x,y), yields a proof of the proposition put-in-the-meter(p(a),p(q(a))). A proof a of the proposition (Σx: person)(Σy: dime)has(x,y), yields, when fully evaluated, a person b, a dime c, and a proof that b has c. The projection p(a) recovers the person b from the proof a, while the projection p(q(a)) recovers the dime c. So a canonical proof of proposition (16) is a function which, when applied to an arbitrary proof that a person has a dime, yields a proof that the person puts the dime in the meter. This assigns too strong a proof condition to sentence (15); if there is a proof that a person has a dime, and a proof that the person does not put that dime in the meter, then there is no proof of (15). This is true even if there are proofs that the person in question put enough dimes in the meter to prevent his car from being ticketed. It can similarly be shown that the proof conditions assigned by Ranta's grammar to sentence (14) are too strong.

Ranta's theory, like classical DRT, yields a strong reading of donkey sentences with relative clauses and the determiner 'every'. Thus, like classical DRT, it gives the wrong result in those cases in which a weak reading of such sentences is the favored one. This may be only a minor problem for the theory; perhaps, as M. Kanazawa has argued (see Kanazawa 1994, pp.122-25) the default reading of donkey sentences with relative clauses and the determiner 'every' is the strong reading, while the weak reading becomes salient only when other factors are in place (for instance, the background assumption that people generally will only put as

much money in a parking meter as is necessary to avoid getting a ticket). But even if it is only a minor problem for the theory, it should not be ignored. At the very least, Ranta should provide an account of the sort of factors which can make the weak reading salient.

A more serious problem facing Ranta's theory is the "proportion problem" (see Kadmon 1987, 1990, Heim 1990 for discussions of this problem), a problem which originally arose within the framework of classical DRT. As noted earlier, classical DRT says that the determiner in sentences such as (9)-(11) is to be interpreted as quantifying over farmer-donkey pairs standing in the owning relation. This interpretation yields the correct truth condiitons for (9) and (10), but not for (11). In order to see this, consider a situation in which there are but two farmers who own donkeys. Suppose that one of these farmers owns just one donkey, and that this farmer does not beat the donkey he owns. Suppose that the other farmer owns a pair of donkeys, and that this farmer beats both of his donkeys (This is the same situation described in Kanazawa 1994, p.111). Classical DRT says that sentence (11) is true in this situation, since the majority of farmer-donkey pairs standing in the owning relation also stand in the beating relation. But sentence (11) is false in this situation, since it is not the case that the majority of donkey-owning farmers beat the donkey(s) they own. The reason that this is dubbed the proportion problem should be clear; whenever there is a donkey sentence with a proportional determiner such as 'most' modifying a noun phrase with a relative clause, then DRT will (often incorrectly) interpret the determiner as quantifying over certain sets of pairs, rather than over individuals.

The proportion problem also arises within Ranta's framework; this shoud be expected, since determiners in sentences such as (9)-(11) are interpreted within this framework as quantifying over pairs of a certain sort (in (9)-(11) these pairs consist of a farmer and a proof that the farmer owns a donkey), not individuals. In order to see how this problem arises within Ranta's framework, consider the following proposition of type theory:

(15) (Most z: (Σx: farmer)(Σy: donkey)own(x,y))beat(p(z),p(q(z))).

This proposition can be derived in an extension of Ranta's grammar, an extension obtained by adding the rules for the quantifier Most given in Göran Sundholm's paper on generalized quantifiers in Martin-Löf's type theory (Sundholm 1989), as well as rules for sugaring propositions containing that quantifier into English sentences. This proposition can then be sugared into sentence (11). The quantifier Most, according to the definition given in Sundholm 1989, counts proofs of the proposition (Σx: farmer)(Σy: donkey)own(x,y); if it is stipulated that there is only one canonical proof that a given farmer a owns a given donkey b, then it counts farmer-donkey pairs which can be proven to stand in the owning relation. But this is the wrong interpretation for sentence (11), since 'most' in that sentence is interpreted as quantifying over donkey-owning

farmers, not farmer-donkey pairs. In particular, the interpretation of 'most' as counting proofs that farmer-donkey pairs stand in the owning relation will get the wrong result in the situation presented above (assuming that the farmer-donkey pairs mentioned in that example can all be proven to stand in the hypothesized relations). This problem will also arise in the analysis of other donkey sentences with proportional determiners and relative clauses, since the determiners in such sentences will be interpreted as quantifying over sets of the form (Σx: A)B(x), whose canonical elements are pairs, not individuals.

The proportion problem is a more serious problem for Ranta's framework than the problem I mentioned earlier in this part of the paper. This is because the "pair-quantificational" (to borrow a term from Kanazawa 1994) reading of donkey sentences with relative clauses and proportional quantifiers is not the default reading of such sentences, whereas the strong reading of donkey sentences with relative clauses and the determiner 'every' may very well be the default reading of *those sentences* . Fortunately, there are a number of possible solutions to the proportion problem which can be implemented in Ranta's framework, two of which I will mention here. The first of these is offered by Sundholm (Sundholm 1989), and makes use of the notion of an "A-injection" (ibid., p.10). An A-injection is a function from a set (Σx: A)B(x) to a set C. It maps elements of the set (Σx: A)B(x) which have distinct first components into distinct elements of C. Since the first components are elements of the set A, the name 'A-injection' is an apt one. Using the notion of an A-injection, the cardinality of the set of first components in a set of the form (Σx: A)B(x) can be defined. One can then define a proof of a proposition of the form (Most z: (Σx: A)B(x))D(z) to be a proof that a majority of the *first components* of the set (Σx: A)B(x) are D. So a proof of (15) would be a proof that a majority of donkey-owners beat the donkey(s) they own. This is in line with the intuitively correct proof conditions for sentence (11). Sundholm himself ultimately rejects this proposed solution to the proportion problem. His reason for doing so is that the definitions of the quantifiers "would no longer be uniform in the domain...of quantification" (Sundholm 1989, p.10). This simply means that the quantifier Most, for example, will sometimes count elements in the domain of quantification (for instance, if the domain of quantification is a simple set such as the set of farmers), and sometimes count elements in a set other than the set forming the domain of quantification (for instance, it counts the first components of the set forming the domain of quantification, if that set is of the form (Σx: A)B(x)). So one would have to define the quantifiers by cases, and this requires, as Sundholm points out, "the principle of definition by induction over the universe of small types" (Sundholm 1989, p.10) in Martin-Löf's type theory. Sundholm prefers not to define the quantifiers in this way, apparently because of the inelegance of the

definition and the strength of the formal machinery necessary for that definition. I think that Sundholm is too hasty here, since the assumption that the quantifiers are uniform in the domain of quantification, although it leads to a more elegant framework for studying generalized quantifiers, also leads to the proportion problem. His suggestion of using A-injections to solve the proportion problem demands more serious attention than he himself gives it.

Another proposed solution to the proportion problem which could be adopted by someone working within Ranta's framework is found in Kadmon 1990. One could argue, as Kadmon does, that donkey sentences such as (11) contain a uniqueness presupposition. That is, such sentences contain a presupposition that a uniqueness condition for the donkey pronoun (in (11), the pronoun 'it') is satisfied. This condition is satisfied just in case each farmer who owns a donkey owns exactly one donkey. If this condition is not satisfied, then such sentences cannot be felicitously uttered in the given context. If this approach is correct, then the proportion problem disappears, since whenever one quantifies over farmer-donkey pairs standing in the owning relation, one is in effect quantifying over donkey-owning farmers. If this approach is adopted to Ranta's framework, then nothing stands in the way of using quantifier definitions uniform in the domain of quantification, as Sundholm prefers. However, one must also provide, within that framework, some account of what occurs when the uniqueness presupposition for a sentence is not met. More importantly, one must meet the objections which have been brought against Kadmon's position (for instance, the arguments in Heim 1990). Despite these obstacles, an approach based on Kadmon's work may be a viable alternative to one based on Sundholm's work.

In conclusion, the data which Ranta cites in support of his contention that his approach to the study of discourse anaphora is superior to the approaches of DRT and DPL do not deliver as clear a verdict as he believes. The anaphoric links which his theory predicts to be present in sentences such as (1), (4), and (6) are not clearly there. Moreover, even if he is correct about the presence of these links, the semantic analysis which he gives of these links may very well be incorrect. On the other hand, many of the readings which Ranta's theory gives to donkey sentences are strikingly similar to readings given by the model-theoretic approaches he criticizes. Indeed, the readings are so similar that problems associated with DRT and DPL are also present in Ranta's theory. Given that this is the case, a proponent of DRT or DPL can look upon Ranta's theory, with some justification, as a theory which is essentially similar to his own, but which in addition predicts a number of questionable anaphoric links which his own theory does not. If a proponent of Ranta's approach is to convince him otherwise, he must provide some defense of the semantic analysis of sentences such as (1) and (4) given in Ranta's approach (in the face of such examples as the situation involving the two donkey-beating farmers given in the first part of this paper), as well as some account of the

difference in acceptability between such sentences as (1) and (6). If this can be done, then this approach to discourse anaphora would justifiably merit the attention of proponents of DRT and DPL.

References

Chierchia, G.: 1990, 'Anaphora and Dynamic Logic', ITLI Prepublication Series LP-90-07, Institute for Language, Logic and Information, University of Amsterdam.

Chierchia, G.: 1992, 'Anaphora and Dynamic Binding', *Linguistics and Philosophy* 15, 111-183.

Heim, I.: 1990, 'E-Type Pronouns and Donkey Anaphora', *Linguistics and Philosophy* 13, 137-177.

Kadmon, N.: 1987, *On Unique and Non-unique Reference and Asymmetric Quantification* , Ph.D. Dissertation, University of Massachusetts, Amherst.

Kadmon, N.: 1990, 'Uniqueness', *Linguistics and Philosophy* 13, 273-324.

Kamp, H.: 1981, 'A Theory of Truth and Semantic Representation', in J. Groenendijk, T.M.V. Janssen, and M. Stokhof (eds.), *Formal Methods in the Study of Language* , Part 1, Mathematical Centre Tracts 135, Mathematisch Centrum, Amsterdam, 277-322.

Kanazawa, M.: 1994, 'Weak vs. Strong Readings of Donkey Sentences and Monotonicity Inference in a Dynamic Setting', *Linguistics and Philosophy* 17, 109-158.

Nordström, B., K. Petersson, and J. Smith: 1990, *Programming in Martin-Löf's Type Theory* , Clarendon Press, Oxford.

Pelletier, F.J., and L.K. Schubert: 1989, 'Generically Speaking, or Using Discourse Representation Theory to Interpret Generics', in G. Chierchia, B.H. Partee, and R. Turner (eds.), *Properties, Types, and Meaning II* , Kluwer, Dordrecht.

Ranta, A.: 1991, 'Intuitionistic Categorial Grammar', *Linguistics and Philosophy* 14, 203-239.

Ranta, A.: 1994, *Type-Theoretical Grammar* , Clarendon Press, Oxford.

Sundholm, G.: 1989, 'Constructive Generalized Quantifiers', *Synthese* 79, 1-12.

Epistemic *must* as evidential

Robert R. Westmoreland
Indiana University

1. Introduction

The English modal auxiliaries *must* and *might*, near-synonyms like *have to* and *may*, and their analogues in other languages such as German *müssen* and *können*, are traditionally analyzed as natural language realizations of the modal-logical necessity and possibility operators n and p. Modal operators are interpreted in a possible-world semantics using an 'accessibility relation' which maps each world in the model onto a set of 'accessible worlds'. A formula of the form n Φ is true in world ω just in case Φ is true in all worlds accessible to ω. Likewise, p Φ is true in world ω just in case there is a world accessible to ω in which Φ is true. Natural language modal expressions like *must* and *might* are identified with n and p respectively, with different accessibility relations yielding the different senses of these words (Kratzer 1976, 1977, 1981, Chierchia and McConnell-Ginet 1991: 234-241).

In natural language modality, the accessibility relation (or 'conversational background' or 'modal base') is not in general made explicit, but is taken from the context. To illustrate, sentences (1) and (2) are to be interpreted as 'deontic' and 'circumstantial' modality, respectively.

(1) Jill must phone her mother
(2) This man must get to the hospital right away (or he will die)

Deontic modality describes a person's duty; (1) might be uttered in a situation where Jill has made a promise, thus incurring an obligation to phone her mother. Circumstantial modality describes the force of circumstances; imagine (2) being said of a man who has suffered a life-threatening injury in an accident. The accessible worlds in the former case, presumably, are those where Jill's obligations are fulfilled; in all such worlds, Jill phones her mother. In the latter, the modal base presumably consists of those worlds where the injured man's life is saved; the claim is that in all these worlds, the man gets to a hospital right away.

An important variety of modality is 'epistemic' modality, illustrated in (3). The accessible worlds that define epistemic modality are those that are consistent with what is known.

(3) (The car is in the driveway, so) Jack must be home

The author rues all shortcomings of this paper that remain despite the invaluable counsel of Alice ter Meulen, Michael Dunn, Paul Dekker, Ariel Cohen, the IU Semantics Reading Group, and the participants of the Tenth Amsterdam Colloquium.

If we know such things as that Jack never goes anywhere without the car, and we learn that the car is in the driveway at Jack's home, then Jack is home in all worlds consistent with our knowledge.

The present paper focuses on *must* in its epistemic sense. The goal is to show that, while the traditional analysis sketched above may be adequate for non-epistemic *must*, it gives fundamentally wrong results in the epistemic case. This is because so-called epistemic *must* is not really epistemic, but EVIDENTIAL. That is, it does not report the speaker's knowledge or information state, but rather the SOURCE of information.

I argue that epistemic *must*[1] should not be seen semantically as a modal at all, despite its categorial status as a 'modal auxiliary verb'. That is, it is not a quantifier over possible worlds, information states, etc. In the following section, I discuss the possible-world analysis of modality in detail, and demonstrate that this approach is problematic for dealing with epistemic *must*. As has been observed (eg., van Benthem 1985), in modal logic $\Box\Phi \rightarrow \Phi$ holds for a reflexive accessibility relation[2], whereas $\text{MUST}\Phi \rightarrow \Phi$ is too strong a claim to make for natural language epistemic *must*. This is because $\text{MUST}\Phi$ seems more readily defeasible than Φ. However, there is a more fundamental problem: if $\text{MUST}\Phi$ is taken to mean that Φ is true in all worlds consistent with speaker-knowledge, then $\text{MUST}\Phi$ becomes for all practical purposes indistinguishable from Φ. Thus, while $\text{MUST}\Phi \leftrightarrow \Phi$ does not strictly hold, the cases where Φ is true and $\text{MUST}\Phi$ false are cases where Φ could not be asserted in actual discourse. Hence, the distinction that this analysis makes between Φ and $\text{MUST}\Phi$ is not the right one.

Perhaps the traditional analysis can be saved by further restricting the accessible worlds to those that are 'normal' in some sense: $\text{MUST}\Phi$ means that Φ is true in all normal worlds consistent with speaker-knowledge. This has the consequence that the accessibility relation is irreflexive (because the actual world may not be normal) and so $\text{MUST}\Phi \rightarrow \Phi$ does not hold. I will argue, however, that the concept of 'normal worlds' cannot be consistently specified, and that the possible-world semantic analysis remains inadequate.

There is a more recent analysis of epistemic modality in terms of dynamic Update Semantics (Beaver 1993, Groenendijk et al. 1994, Veltman to appear), examined in §3. In this framework, the semantic value of a sentence is conceived to be the information which the sentence adds to a context, where a 'context' is some sort of database. This information is called the sentence's 'Context Change Potential' (CCP, after Kartunnen 1973 and Heim 1982). Advocates of this approach hold that sentences with epistemic modals have a vacuous CCP; they are thus

1 I will continue to use the established term 'epistemic *must*', though my argument is that the term is inaccurate.

2 To say that the epistemic accessibility relation is irreflexive would be tantamount to supposing that the world might be inconsistent with our knowledge, which is clearly absurd.

thought of not as expressing propositions which are true or false, but as constituting 'tests' which adjudge a statement to be 'acceptable' or not. A formula of the form $\text{MIGHT}\Phi$ is acceptable in a context iff updating the context with Φ would not lead to absurdity—ie, Φ is compatible with the database. A formula of the form $\text{MUST}\Phi$ is acceptable in a context iff updating the context with Φ would not change the context—ie, Φ is entailed by the database. However, the Update Semantic account also fails to correctly distinguish $\text{MUST}\Phi$ from Φ, and for essentially the same reason; both $\text{MUST}\Phi$ and Φ ultimately provide the listener with the same information: that the speaker considers Φ to be true.

Both possible-world semantics and Update Semantics treat *must* and *might* as duals; that is, they assume that in some sense $\text{MUST}\Phi = \neg\text{MIGHT}\neg\Phi$ and $\text{MIGHT}\Phi = \neg\text{MUST}\neg\Phi$. In so doing, they miss the key aspect of the meaning of *must*. As suggested in example (3) above, $\text{MUST}\Phi$ relates Φ to some other information that serves as evidence for Φ. The speaker of (3) does not know that Jack is home; rather, the speaker knows that Jack's car is in the driveway, and infers Jack's presence on the basis of this. The use of *must* labels a proposition as an inference. Modeling *must* as a quantifier over possibilities does not capture this crucial fact. This is discussed in section §4. A linguistic device that does not depend on the information which the speaker has or wishes to convey, but rather on the way in which the speaker arrived at that information, is called an evidential. Evidentials are quite common cross-linguistically, so it is important to carefully examine the rôle they play in the composition of sentence meaning. An evidential should not be seen as a quantifier over propositions, or as a function from propositions to propositions, but as something like a 'label', providing information about the proposition. In a sense, it functions at a metalinguistic or meta-logical level with respect to the propositional content.

Because it gives information about, rather than adding to or modifying, the proposition over which it has scope, an evidential like *must* allows the hearer to evaluate the reliability of the proposition. On one hand, this is information that would traditionally be labeled pragmatic rather than semantic; yet on the other hand it is part of the sentence meaning, independent of speaker meaning. In section §5, I re-examine the semantics/pragmatics distinction as it is commonly formulated, suggesting that there is an important overlap, a portion of the semantics that is pragmatic in nature. I propose that a compositional semantics of sentence meaning involves two independent levels composed simultaneously: in addition to the propositional content there is a level of metalinguistic information. This relates to the analysis not only of evidentials but of sentence adverbs, questions, and other phenomena.

2. The possible-world semantic analysis

Here, I present the formal apparatus needed for the standard possible-world semantic analysis of *must*, essentially following Kratzer 1981.

- A set W of worlds.
- The set P of (atomic) propositions $\{\Phi | \Phi \subseteq W\}$.
- A set P^+ of atomic propositions plus non-atomic propositions, of which we only need consider the modal propositions MUST Φ and MIGHT Φ for all $\Phi \in P$.
- The set R of accessibility relations $\{r | r \in W \times W\}$. It is customary in modal logic to deal with an accessibility relation, but for our purposes it will be more convenient to deal with sets of accessible worlds. For each $r_i \in R$ define a function f_i from worlds to sets of worlds such that for each world ω, $f_i(\omega) =_{df} \{\omega' | r_i(\omega, \omega')\}$. Such a set $B = \{\omega' | r_i(\omega, \omega')\}$ is a modal base. (Note that a modal base, being a set of worlds, is also a proposition.)
- An atomic proposition Φ is true at world ω iff $\omega \in \Phi$.
  A modal proposition MUST Φ is true at world ω with respect to a modal base B iff $\forall \omega' \in B. \omega' \in \Phi$.
  A modal proposition MIGHT Φ is true in a world ω with respect to a modal base B iff $\exists \omega' \in B. \omega' \in \Phi$.

The 'worlds' in this model correspond to possibilities, and the 'propositions' are supposed to model the propositional content of natural language declarative sentences thus: the content of "Jack and Jill went up a hill" corresponds to that set of worlds where Jack and Jill did, in fact, go up a hill.

The modal bases give us different senses of *must* and *might*. Let us reconsider examples (1) - (3), repeated below.

(1) Jill must phone her mother
(2) This man must get to the hospital right away
(3) Jack must be home

In principle, a modal base can be any set of worlds, but in practice the modal bases needed for natural language semantics would seem to be expressible by ordinary sentences. So, an appropriate modal base for (1) is that set of worlds where "Jill fulfills her obligations" is true; for (2), those worlds where "This man survives" is true; for (3), those where "What the speaker knows is true" is true.

Informally stated, the modal base for epistemic modality is the set of worlds consistent with the speaker's knowledge. It is part of what we mean by 'knowledge'—the 'factivity' of knowledge—that what is known cannot be inconsistent with the world of evaluation (the 'real world'). On this basis, then, MUST $\Phi \rightarrow \Phi$. Yet intuitively, MUST Φ makes a weaker

claim than Φ. If I utter, "Jack is home" and it turns out that Jack is in fact not home, then I have lied, or at least spoken too rashly. Not so if I have merely said, "Jack must be home" (sentence (3)). The sentence with *must* leaves a lingering scent of doubt. On the other hand, if I know that Jack is not home, or if I have no particular reason to suspect that he is home, then to utter (3) would decidedly be dishonest. Thus the use of MUST Φ is warranted, not by the speaker's knowledge of Φ, but by the existence of a good reason to believe Φ and the lack of a reason to doubt it.

Consider possible contexts for example (3). We know that Jack's car is at his home, but we don't absolutely know that Jack is home. That he has gone out on foot, or is riding with someone else, or has bought a second car, or has been kidnapped, etc., is logically possible, however wildly implausible any of these possibilities may be, given Jack's habits and the nature of the world. To get the logic right, then, the epistemic base should apparently be restricted somehow to 'normal' or 'plausible' worlds. Intuitively, (3) seems to mean that Jack is home in all normal worlds.

Kratzer (1981: 45) proposes a 'stereotypical conversational background' to provide an 'ordering source' for a logic of degrees of possibility. She also suggests it as a potential modal base for *must*[3]. Perhaps we can press it into service here as a formalization of the notion of 'normality'.

> A stereotypical conversational background is a function f which assigns sets of propositions to members of W such that for any w ∈ W:
> f(w) contains all those propositions p such that it is the normal course of events in w that p – for someone, for a community etc.

(Her function f defines a modal base, like our function f_i above)

Veltman (to appear) has a similar notion of 'normal world' in his discussion of modality, although he does not specifically address the semantics of *must*. His definition is rather complicated, and it would take us rather far afield to go through it in detail here, but it crucially depends on a set "of all propositions that a certain agent considers to be normally the case".

Unfortunately, both of these definitions are vague in a crucial way: for what does it mean for a proposition to be "the normal course of events" or "normally the case"? Most propositions do not express a COURSE of events at all, but rather simply an event, or a state. A sentence that reports a

3 She defines a number of modal bases (pp. 44-5), then says that any of them may be the modal base for the German verb *muß* (p. 45). Her primary purpose for introducing the stereotypical conversational background, though, is to account for expressions of comparative possibility like *more likely* and *a slight possibility*. In fact, she claims that epistemic *muß* has the epistemic conversational background, the set of worlds consistent with speaker knowledge. This is the modal base which we have rejected as too inclusive.

course of events is at least bi-clausal, because it must report at least two events: eg, "Jill got up and immediately brushed her teeth". A proposition that expresses a normal course of events would have to be a generic conditional: "When Jill gets up she immediately brushes her teeth". Similar considerations apply to Veltman's notion. Very few affirmative propositions are normally the case; for example, the sentence "Jill is brushing her teeth" is false in most cases, while "Jill is not brushing her teeth" is normally true. Do we want the least exceptional worlds to be those in which Jill never brushes her teeth? Actually, it seems clear from the sentence schemata he discusses ("P's normally are R") and from his concluding remarks that Veltman has generic conditionals in mind as well, rules which characterize an agent's understanding of the world.

Veltman makes the remarkable claim that there should be at least one world in which every normal proposition holds. "If it is not even conceivable that everything is normal, something is wrong." Similarly, Kratzer (p. 47) describes such a world "in which the normal course of events is realized" as "a complete bore. There are no adventures or surprises." Yet something IS wrong, for it is surely not conceivable that EVERYTHING is normal. The reason is simply that rules normally have exceptions; in virtue of this fact, a world where no rules had exceptions would be a world where all rules had exceptions.

There is nothing paradoxical or irrational about expecting the unexpected, about saying that it is normal for the abnormal to take place. This is why people buy insurance policies and lottery tickets. But even if we can somehow exclude meta-rules from our definition of normality, so that statements like "Rules normally have exceptions" do not apply, there can still be no completely normal worlds. First, there may be a set of generalizations such that each one is normally realized, yet the entire set is normally not realized. Suppose I have a routine schedule for Mondays, but also most Mondays my wife phones at some unpredictable time and has me drive her to the store. For each scheduled activity, then, I expect to be engaged in the appointed activity at the appointed time, yet I also expect to depart from my schedule at some (unknown) time during the day. Given that my routine is routinely interrupted, there can be no completely normal worlds.

Second, there are rules that govern exceptions. Hence, a certain state of affairs can be normal from one perspective and exceptional from another. Consider the famous generalization about English orthography:

(4) If a word contains the letters E and I in succession, normally the I precedes the E, unless the pair directly follow a C, in which case the E normally precedes the I

Suppose you ask me how to spell the word 'concierge'. I may reply as in (5) or as in (6).

(5) The rule is I before E so it must be C-O-N-C-I-E-R-G-E
(6) The rule is I before E except after C so it must be C-O-N-C-E-I-R-G-E

In setting up our completely normal worlds, should we count the proviso "except after C" as part of the rule, or as an exception? How is 'concierge' spelled in a world without adventures or surprises? The decision—where to draw the line between rules and exceptions—would seem to be arbitrary.

Third, some exceptions are absolute; they are exceptional in all consistent worlds. For example, prime numbers normally are odd. Yet there are no completely normal worlds where the number 2 is odd, composite, or non-existent. Or consider the fact that normally (ie, on the basis of mathematicians' experience), an even number greater than 4 is expressible as the sum of two odd primes. This is the unproved Goldbach Conjecture:

(7) Presumably, every even number greater than 4 is the sum of two odd primes

Goldbach's conjecture is correct or incorrect in all worlds, independently of what is normal.

It is impossible, then, to model epistemic *must* as a quantifier over completely normal worlds, for there are no such worlds. However, it is not necessary that everything be normal for a sentence such as (3) to be reliable. So long as Jack has not departed from his habits, has not been kidnapped, etc., we may reasonably suppose that he is home, regardless of whether there are three-legged dogs, weirdly-spelled words, or prime numbers divisible by 2. In fact, there is really only one generalization that is relevant to (3):

(8) Wherever Jack's car is, Jack is nearby

In all worlds where this generalization in exceptionless, Jack is home. Similarly for (5): in all worlds where I invariably precedes E, 'concierge' is spelled C-O-N-C-I-E-R-G-E. So the correct semantics for epistemic *must* would seem to be this: MUST Φ is true iff Φ is true in all worlds consistent with the speaker's knowledge such that a certain (contextually given) generalization is exceptionless.

- $B = \{\omega' | r_i(\omega,\omega')\}$ for $r_i \in R$ is an epistemic modal base for the speaker iff for all $r_i(\omega,\omega')$:
  (i) there are no propositions Γ and Δ such that the speakers knows Γ, $\Gamma \rightarrow \neg\Delta$, and $\Delta \in \omega'$ (ie, accessible worlds are consistent with speaker knowledge);
  (ii) there are propositions Ψ and Φ such that the speaker knows Ψ, the speaker believes that generally Φ whenever Ψ, and $\Phi \in \omega'$ if $\Psi \in \omega'$

We have gone a long way to 'save the phenomenon', that is, to analyze epistemic *must* as a quantifier over possibilities, but the present proposal still does not yield the desired results, for conditions (i) and (ii) above may interact in such a way as to yield an empty modal base.

Suppose the speaker knows that Jack suffered a stroke and was taken to the hospital this morning in an ambulance, but wishes to deceive the hearer, so utters (3). In this case, we would not want to say that (3) is true. Yet, there are no worlds consistent with the speaker's knowledge where generalization (8) is exceptionless; consequently it is trivially true that in all such worlds, Jack is home. Hence, (3) would be true on the present analysis.

The attempt to model epistemic *must* as universal quantification over possible worlds, then, fails. Quantification over worlds consistent with speaker knowledge is too strong, yielding the undesired entailment $\text{MUST}\,\Phi \rightarrow \Phi$. Further restricting the modal base to completely normal worlds is not possible simply because such worlds are inconsistent, while restricting it to worlds that are normal in some contextually given respect yields undesired results when the modal base is empty.

What is important in claiming $\text{MUST}\,\Phi$ is that the speaker have reason to believe that Φ and no reason to believe otherwise. If anything, *must* is existential rather than universal. Suppose the context contains a set P of propositions and a set G of generalizations, that is, a set of pairs $\langle \Psi, \Phi \rangle$, $\Psi, \Phi \in P$, such that generally Φ whenever Ψ. Then a speaker may state $\text{MUST}\,\Phi$ iff both $\exists \Psi \in P. \langle \Psi, \Phi \rangle \in G$ and $\neg \exists \Psi \in P. (\langle \Psi, \neg \Phi \rangle \in G)$. The latter formula guarantees consistency with speaker knowledge, but it is merely a rider to the former, which is the real gist of *must*, and is not taken into account by the possible-world semantic analysis.

3. The Update Semantic analysis

A difference between non-epistemic and epistemic modality that has been glossed over in the discussion so far is that the latter but not the former depends on the information available to the speaker. Thus, to say that $\text{MUST}\,\Phi$ is true with an epistemic sense is to use the word 'true' in a somewhat unnatural way, for it can only be true relative to a speaker in a situation. Ordinarily we consider a proposition to be simply true or false, and not true-for-A-and-false-for-B. If we want our formal notion of truth to correspond to this intuitive notion, then, it is not appropriate to define truth conditions for sentences with epistemic modals.

Note that in linguistic environments where propositional content is at issue, the use of *must* is awkward at best. This is because $\text{MUST}\,\Phi$ does not express a proposition (ie, it is not simply true or false). It does not occur in factive environments, (9), because facts are simply true or false; it is not used in alternative questions, (10), because truth of a proposition is what is at issue in such a question; and it cannot be part of backgrounded information, (11), because it is propositional content that is presupposed.

(9) I believe / *?know it must be raining
(10) *?Must it be raining?
(11) It must be the butler that did it / *?It's the butler that must have done it

That epistemic modals do not express propositions is acknowledged in dynamic semantic frameworks such as Update Semantics, where "the notion of truth ... loses its key role" (Groenendijk, et al. 1994: 2). Instead, the semantic value of a sentence is conceived to be its CCP, the change that it can bring about when introduced into a context. Epistemic modals, on this view, do not update a context, but rather provide information about the context itself.

For an Update Semantic account of epistemic modality, we may use some of the same formal apparatus that we used above (essentially following Beaver 1993):

- A set W of worlds.
- The set P of (atomic) propositions $\{\Phi | \Phi \subseteq W\}$.
- A set P^+ of atomic propositions plus non-atomic propositions, of which we only need consider the modal propositions $\text{MUST}\,\Phi$ and $\text{MIGHT}\,\Phi$ for all $\Phi \in P$.

In addition, we need contexts and an update function. For present purposes, we may model contexts and atomic propositions both as sets of worlds:

- The set S of contexts (or information states) $\{\Phi | \Phi \subseteq W\}$
  The empty context is written $\perp$
- A function from $S \times P^+$ to S
  We write $\sigma[\Phi] = \tau$, where $\sigma, \tau \in S$ and $\Phi \in P^+$
- If $\sigma[\Phi] \neq \perp$, then $\sigma[\text{MIGHT}\,\Phi] = \sigma$, otherwise $\sigma[\text{MIGHT}\,\Phi] = \perp$
- If $\sigma[\Phi] = \sigma$ then $\sigma[\text{MUST}\,\Phi] = \sigma$, otherwise $\sigma[\text{MUST}\,\Phi] = \perp$

On this account, *must* and *might* are informationally redundant in a well-formed discourse (ie, one which does not lead to the absurd state $\perp$). A virtue of such a semantics is that it accounts for why (12) is well-formed but (13) is not.

(12) Who's at the door? ... It might be Jack ... It's Jill
(13) *Who's at the door? ... It's Jill ... It might be Jack

If epistemic modals are semantically vacuous, conveying no information (having no CCP), what purpose do they serve? The answer is that they do convey information, but at a metalinguistic or meta-logical level; rather than participating in the system whereby propositions update contexts, they provide information ABOUT propositions and their compatibility with contexts. Veltman writes concerning *might* (pp. 7-9):

> The idea behind the analysis of 'might' is this: One has to agree to *might*ϕ if ϕ is consistent with ones knowledge .... Otherwise, *might*ϕ is to be rejected. ... If ϕ is acceptable in your information state σ, you must accept *might*ϕ. And if ϕ is not acceptable in σ, neither is *might*ϕ. Clearly, then, sentences of the form *might*ϕ provide an invitation to perform a test on σ rather than to incorporate some new information in it.

A blurring of the semantics/pragmatics distinction can be detected in this passage, for the semantic value of MIGHT Φ is being described, not in terms of the logical inferences that it licenses, but in terms of how a rational agent behaves in the face of MIGHT Φ, accepting it or not. Epistemic-modal sentences are metalinguistic with respect to non-epistemic-modal sentences, testing them for acceptability in a context. Presumably, such non-informative tests are useful in natural discourse to confirm that speaker's and hearers' information states are compatible. Should they prove incompatible in some respect, then rational agents will revise their contexts accordingly. A full Update Semantic account of epistemic modality, then, ought to deal with the updates (or revisions) licensed by the metalinguistic information conveyed.

A notable characteristic of natural language is that it can be used as its own metalanguage; not surprisingly, then, it is possible to mimic the effect of *might* by making an overt statement about the context. Thus, (14) and (15) parallel (12) and (13).

(14) Who's at the door? ... I don't know if it's Jack or not ... It's Jill
(15) *Who's at the door? ... It's Jill ... I don't know if it's Jack or not

(In fact in Japanese, which has no word for *might*, the closest equivalent is an expression *ka mo shirenai*[4], literally "can't tell whether or not".) Also, (16) and (17)—which are well-formed in contrast to (13) and (15)—are parallel.

(16) Who's at the door? ... It's Jill ... It might have been Jack
(17) Who's at the door? ... It's Jill ... I didn't know if it was Jack or not

However successful this analysis may be for *might*, though, it is not adequate for *must*. To be sure, informationally redundant utterances are not uncommon in discourse; they are used to confirm that speaker and hearer are operating from compatible contexts (Walker 1992). However, it is decidedly odd to use *must* for such purposes, as (18) illustrates.

(18) A: [Question] Who's at the door?
B: [Answer] It's Jill.
A: [Confirmation Request] It's Jill?
B: [Confirmation] It's Jill. / *It must be Jill.

4 Sometimes abbreviated to *kamo* or *kashira*.

The Update Semantic account wrongly predicts that both (19) and (20) are well-formed.

(19) It's Jill ... It's Jill
(20) *It's Jill ... It must be Jill

This is because it fails to genuinely distinguish Φ from MUST Φ. Suppose speaker A utters either Φ or MUST Φ. There are three cases. If Φ is incompatible with hearer B's information state or if $\neg\Phi$ is incompatible with hearer B's state, then B's state is not updated, regardless of whether A has said Φ or MUST Φ. Update Semantics predicts a difference only in the third case, when both Φ and $\neg\Phi$ are compatible with B's state. If A has uttered Φ and B accepts it (which he will presumably do if he believes that A's information is reliable), then B updates his state with Φ; if A has uttered MUST Φ, on the other hand, he does not. However, MUST Φ provides metalinguistic information about Φ; in particular, the fact that A has said MUST Φ indicates that A has the information that Φ. Consequently, it is reasonable for B to revise his state to accommodate Φ, which he presumably will do if he believes that A's information is reliable. So although Φ and MUST Φ have different context UPDATE potentials, the resulting context CHANGE is the same. The change occurs due to the logic when Φ is uttered, meta-logically when MUST Φ is uttered, but this difference is theory-internal. Update Semantics does not predict an empirical difference between Φ and MUST Φ.

Thus, both possible-world semantics and Update Semantics fail to properly distinguish Φ and MUST Φ. In their non-epistemic uses, *must* and *may /might* are duals. So if (21) is true then so is (22), and vice-versa.

(21) It is not the case that Jill may skip the meeting
(22) Jill must attend the meeting

We are thus seduced into assuming that they are duals in their epistemic uses as well. However, (23) is simply equivalent to (24). Therefore, the common approach of defining *might* then taking *must* to be its dual results in a semantics for MUST Φ that is really appropriate for Φ.

(23) It is not the case that Jill might be absent from the meeting
(24) Jill is present at the meeting

If we examine epistemic *must* and *might* closely, we will see that their behavior is different with respect to tense. First, unlike *must*, *might* can refer to future possibilities ((25) and (26)). Further, *might* but not *must* can be used from the perspective of a previous context ((27) and (28)).

(25) It might rain tomorrow
(26) *It must rain tomorrow
(27) As far as we knew, Jack might have been home, but he isn't
(28) *Given our assumptions, Jack must have been home, but he isn't

Thus their semantic relationship to one another cannot be as simple as being duals.

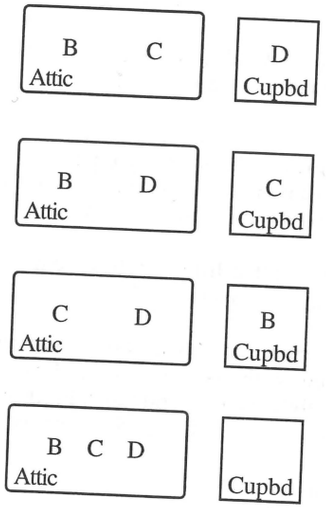
4. *Must* as evidential

Previous accounts have the correct intuition that *must* has something to do with what is normally the case, but the device of quantifying over normal worlds does not yield the right logic. Update Semantics rightly intuits that MUST Φ does not express a proposition within the discourse, but rather provides meta-logical information about the relation between the speaker and the informational content of Φ. What existing accounts lack is a sense of what that relation is. A speaker uses *must* to make an assertion based on a deduction rather than on immediate knowledge.

Beaver (1993: 8-9), illustrates his Update Semantics account of *must* and *might* with this story:

- Anna is seeking Bertha, Clothilde and Daisy ...
- Anna ... knows that the only remaining hiding places are the cupboard (which is not in the attic) and the attic (which is not in the cupboard.)
- Only one person fits in the cupboard.

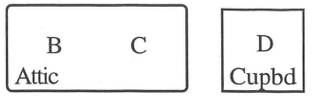
At this point, Anna's information state contains four possible worlds (B=Bertha, C=Clothilde, D=Daisy):



On this basis, Anna is entitled to utter (29) (cf. Beaver's E16).

(29) Bertha might be in the cupboard

Suppose Anna opens the door and finds Daisy in the cupboard. Her epistemic state now is reduced to one world:



On this basis, Anna can utter (30). But she cannot utter (31) in this situation; rather she would say (32).

(30) Bertha must be in the attic
(31) *Daisy must be in the cupboard
(32) Daisy is in the cupboard

There is no reason why (31) should be disallowed if the use of *must* depends simply on the context. On the possible-world semantic view, (31) is true; on the Update Semantic view, it is acceptable. Why, then, is (31) infelicitous? The difference between the information in (30) and (32) lies not in the context they are drawn from, but in the way that context was created. Anna knows that Daisy is in the cupboard because she sees her there; this information is incontrovertible. On the other hand, that Bertha is in the attic is an inference. Suppose now that Anna runs to the attic and finds, to her surprise and ours, no one there. Unbeknownst to Anna, the other girls have grown tired of the game and have gone off to Clothilde's home to watch television. You may object that this is not possible given the way the situation was defined. But the initially-given conditions were in fact quite unrealistic. In real life, we can rarely if ever be absolutely certain of circumstances that are not actually present to our senses. Only mathematicians and semanticists have this luxury.

The essential aspect of the meaning of epistemic *must* is that it labels a proposition as being inferred rather than known. This aspect is not captured in the traditional semantic analyses. It is instructive to consider some examples from actual texts[5]. (Underlines added.)

(33) If all one knows is
premise 1: Q's normally are not R
premise 2: P's normally are R
premise 3: x is P
premise 4: x is Q
then it remains open whether one can presume that x is R. Clearly the object x must be an exception to one of the rules, but there is

5 (33) is from Veltman 1991, p. 5. (34) is from Arthur Conan Doyle's story "The Sign of Four". (35) is from Robert Hetzron, "Afroasiatic Languages" in Bernard Comrie (ed.) (1990) *The World's Major Languages* p. 650. (36) is from Lewis Carroll's *Alice's Adventures in Wonderland.*

no reason to expect it to be an exception to the one rule rather than to the other.

(34) Why, of course I knew that you had not written a letter, since I sat opposite to you all morning. I see also in your open desk there that you have a sheet of stamps and a thick bundle of postcards. What could you go into the post-office for, then, but to send a wire? Eliminate all other factors, and the one which remains must be the truth.

(35) There is evidence for several lateral consonants in Proto-Semitic; they are still used in modern South Arabian, South Cushitic and some Chadic languages (e.g. *balsam* ultimately comes from the Semitic root *bśm* where *ś* must have been a lateral fricative).

(36) So she began. "O Mouse, do you know the way out of this pool? I am very tired of swimming about here, O Mouse" (Alice thought this must be the right way of speaking to a mouse; she had never done such a thing before, but she remembered having seen, in her brother's Latin Grammar, "A mouse—of a mouse—to a mouse—a mouse—O mouse!")

In all these cases (and in all cases that I have observed), the statement made with *must* reports a conclusion drawn from some evidence which is part of the context. This is clear in (33), which takes the form of a classical syllogism. If the premises hold, then the object x is an exception to one of the rules. It is cases such as this which suggest that *must* is the logical necessity operator.

However, outside of mathematical and logical discourse, the use of *must* to mark a logical deduction is rare. More commonly, it marks the conclusion of an informal 'common-sense' deduction or of an abduction. This is exemplified in (34). Sherlock Holmes is explaining how he deduced that Watson's trip to the post office had been taken in order to send a telegram. As a rule, a person goes to the post office to mail a letter, to purchase stamps or post cards, or to send a telegram. Holmes' uncanny powers of observation allow him to eliminate the first two factors and conclude that Watson's trip had the third purpose. It is easy enough to pick holes in the great detective's reasoning: perhaps Watson had written a letter yesterday, or perhaps he had arranged to meet someone at the post office, etc. Cases such as this seem to be the motivation for the restriction to 'normal' worlds, for Holmes here makes tacit assumptions that would hold "in the normal course of events".

We have seen that assumptions like this can be treated as a binary relation on propositions: whenever Ψ holds generally Φ holds as well. Holmes reaches his conclusion by reasoning from principles like this. The limiting case, when the relation between Ψ and Φ holds not generally but always, is a logical deduction. But what such principles could be behind the reasoning in (35)? It is a principle of historical linguistics that whenever a string of phonemes β derives from another string α, there is

some systematic relationship between β and α. In particular, when a language borrows a word from another language, the word is adapted to the phonotactics of the borrowing language. But these principles alone do not lead ineluctably to the conclusion from ls<ś that [+lateral,+fricative] ∈ ś. Rather, someone had the creative insight that this was a reasonable explanation; this plus the fact that no other possible explanation suggests itself leads to the sense that it 'must' be the case. Finally, the connection between the evidence (the Latin Grammar) and the conclusion in (36) is entirely unprincipled; Alice's reasoning is based on a childish hunch. So there is an extent to which our reasoning is principled, and can be modeled with some kind of set-theoretic logic, but there is an extent to which our reasoning is creative, based on hunches and feelings whose provenance is obscure. In either case, we are entitled to use *must* to present our conclusions. Therefore, the semantics of *must* cannot be formalized by simply incorporating a common-sense logic.

It is not unusual for languages to mark the speaker's source of information. A morpheme which has this function is called an 'evidential'. Evidentials are quite common cross-linguistically, though not surprisingly they play a greater rôle in some languages than in others. In addition to items like *must*, which label an inference, evidentials common to many languages label a proposition as being what the speaker has been told, as what the speaker has experienced, or as a speculation. Less common evidentials may label a proposition as based on visual evidence, non-visual sensory evidence, etc. In the Siouan language Hidatsa, for example, there is a set of five clause-final morphemes that indicate that a proposition is, respectively, (i) true, (ii) believed to be true, (iii) taken to be common knowledge, (iv) what the speaker has been told, or (v) a speculation; in addition there is a sixth morpheme which makes the sentence a question (Palmer 1986: 70, citing Matthews 1965).

Evidentials do not make assertions. That is, the propositional content (the truth value or CCP) of EVIDΦ, where EVID is an evidential, is simply the content of Φ. Thus MUST Φ is not uninformative, as Update Semantics would suggest, but rather doubly informative: it contributes the information TO the context that Φ, and the information ABOUT the context that Φ is an inference. In the unmarked case, Φ is asserted, that is, presented as knowledge. With an evidential, Φ is presented as inference, hearsay, speculation, etc.

In this regard, it is noteworthy that the Hidatsa question-marker appears to belong to the same morphosyntactic system as the evidential markers. No one would suppose that a question asserts a proposition. Rather, a question is a kind of metalinguistic operation on a proposition, presenting it as one whose truth value needs to be ascertained. Likewise, an evidential sentence does not assert a proposition either. A sentence with a 'hearsay' marker presents a proposition as being second-hand information; a sentence with *must* presents a proposition as an inference; and so on.

Evidentials bear an interesting relation to the 'Defeasibility Hierarchy', given in (37), that has been proposed to account for certain discourse phenomena (Walker 1992; cf. Galliers 1991).

(37) hypothesis < default < inference < linguistic < physical

The hierarchy reflects the relative defeasibility of information in a context, based on the source of that information. Thus, speculation is overruled by default assumptions, which can be overruled by inference, which is overrulable by a reliable report, which in turn falls to physical (sensory) evidence. Notably, it is precisely this kind of information that is grammaticized in evidential markers[6]. English marks inference (*must*) and hypothesis (*might*). This is why Anna says (32) instead of (31), and why Φ and MUST Φ are not the same.

5. Semantics or pragmatics? Redrawing the line

Jack proposes eating out tonight, and Jill volunteers (38):

(38) Pete's Diner is a good restaurant

Jack understands from this that Jill would like to eat at Pete's Diner.

There are two steps to arriving at this understanding, involving what are known respectively as semantics and pragmatics. Semantics tells Jack that the world which he inhabits is such that the entity identified as 'Pete's Diner' belongs to the set of 'good restaurants'. Semantics deals with sentence meaning, which is relatively independent of context[7], and is based on a theory of logical entailment. That is, one may reason deductively using the sentence meaning and the principles of logic. Given that (38) is true in world ω, for instance, one knows ineluctably that (39) is true in world ω.

(39) Pete's Diner is a restaurant

Semantics is compositional: given the meanings of lexical items and a phrase marker, the sentence meaning can be derived mechanically.

On the other hand, the inference that Jill would like to eat at Pete's Diner is pragmatic. Pragmatics tells us what the speaker means by an utterance. Pragmatic interpretation depends crucially on context, and is based on a theory of rational behavior. That is, one must consider what goal the contribution of a given sentence meaning would rationally serve

6 Default assumption does not seem to be as commonly marked as the others, although the Hidatsa marker of "common knowledge" might be considered to be such. Also, physical evidence is usually the unmarked case, although there are languages that have overt markers for this.

7 This is not to say that context is irrelevant to semantic interpretation. For instance, deictic items draw their denotation from context. But one does not have to know anything about the context to know that the word "I" refers to the speaker in that context. By contrast, one cannot tell anything about the speakers intention in uttering "I'm cold" without some understanding of the context in which it is uttered. This is the distinction I mean by 'context dependent vs independent'.

in the given context. Pragmatic reasoning is heuristic and defeasible. For example, perhaps Jill does not in fact really care about eating at Pete's Diner, but she believes that Jack would enjoy it, and has actually uttered (38) motivated by a wish to please Jack.

Pragmatics is, in fact, not even particularly linguistic. Any human behavior—a laugh, a gesture, the clothes a person has chosen to wear—is interpreted by the same principles[8]. The study of pragmatics embraces the way we interpret one another's behavior in terms of context and goals; language just happens to be an especially conspicuous form of human behavior. Natural language semantics, in contrast, seeks those principles of compositionality that are proper to human language. The distinction between semantics and pragmatics is clear-cut, and it is important not to muddle them.

(40)

Semantics	Pragmatics
Sentence Meaning ...	... Speaker Meaning
Context-Independent ...	... Context-Bound
Compositional ...	... Heuristic
Specific to Language ...	... General Behavior
Logical Entailment ...	... Rational Behavior

So in claiming that epistemic *must* does not contribute to propositional content, but rather shows something of the speaker's intention, am I not making this fatal conceptual error? Have I not confused the semantics of *must* with its pragmatics?

The meaning of epistemic *must* seems to straddle the fence. It is pragmatic-like in that its use pertains to a theory of rational behavior rather than one of logical entailment. One may not reason with MUST Φ because MUST Φ is not a proposition. Rather, given MUST Φ, one has to decide, based on heuristic, common-sense principles, whether or not to reason with, and act on the basis of, Φ. This decision seems to depend on the quality of the evidence on which Φ is based, the difficulty of verifying Φ directly, and the consequences of acting on the basis of Φ if it happens not to be correct after all. Anna, concluding that Bertha must be in the attic, will then go to the attic to see with her own eyes. Poor Alice, concluding that 'O mouse' must be the proper way to address a mouse, and lacking the means to obtain more dependable information, is forced to assume that it IS the proper locution, act accordingly, and hope for the best.

On the other hand, *must* is semantic-like in that it contributes to sentence-meaning in a context-independent way; namely, it contributes the information that the propositional content of the sentence is inferred rather than known. This is context-independent in that, encountering a sentence like (3) 'out-of-the-blue', divorced of context, you would know

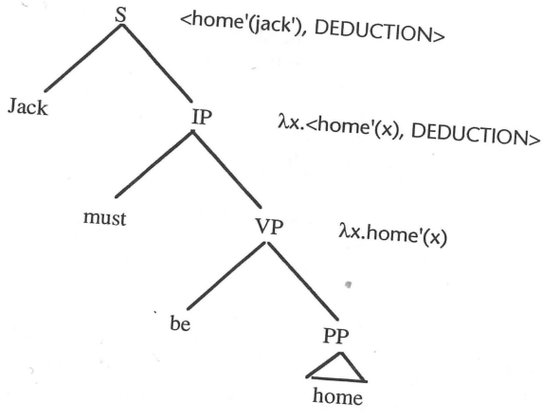
8 Pragmatics may not even be particularly human. Don't we interpret animal behavior along much the same lines?

that it expressed the speaker's inference. (In contrast, you would not know genuinely context-dependent information like who the speaker is or what goals the speaker sought to accomplish by uttering (3).) Moreover, this contribution is made compositionally. The semantics of a sentence can be thought of as an ordered pair $\langle \Phi, \text{LABEL} \rangle$, where Φ is the propositional content and LABEL provides information regarding Φ, such as the speaker's source of information. Suppose the syntactic category of *must* is $(\text{NP} \setminus \text{S}) / \text{VP}$ and that NP and VP are semantically of types e and $\langle e,t \rangle$ respectively. Then *must* has the semantics given in (41):

(41) $\lambda f \lambda x \langle f(x), \text{DEDUCTION} \rangle$

The composition proceeds as in (42):

(42)



Of course, this kind of analysis pertains not just to one sense of one English word *must*, but to the multitude of evidential markers found cross-linguistically. That is, it is a non-trivial discovery about the human language faculty that closed-class lexical items (contrary perhaps to our wishes or expectations) sometimes have this kind of speaker-dependent, metalinguistic interpretation, and a formal theory of natural language semantics will have to expand to make room for this fact.

Consequently, a more realistic conception of the semantics/pragmatics distinction, as it applies to natural language, is that suggested in (43). There is a portion of pragmatics that overlaps with semantics.

(43)

Semantics	**Pragmatics**
Sentence Meaning ...	... Speaker Meaning
Context-Independent ...	... Context-Bound
Compositional ...	... Heuristic
Specific to Language ...	... General Behavior

Logic	**Meta-Logic**
Logical Entailment ...	... Rational Behavior

Actually, there are many linguistic expressions that do not add to propositional content, but rather label that content in some way. For example, there is no truth-conditional (or CCP) difference between (44) and (45).

(44) Jack is at the office
(45) Actually, Jack is at the office

The difference is that, in addition to making an assertion about Jack's whereabouts, (45) labels the assertion as potentially coming as a surprise to its addressee. Sentence adverbs like *actually*, *frankly*, and *hopefully* belong to the metalinguistic semantic-pragmatic overlap, as do the sentence-final particles in, for example, Mandarin or Japanese, that urge an assertion on the hearer, seek the hearer's acquiescence, etc. The conjunction *but* does not differ from *and* in its contribution to truth conditions or CCP, but provides information about how the conjoined propositions relate to the context.

Even questions might best be analyzed as labeled propositions. (Recall that the Hidatsa question marker belongs to the same syntactic category as the evidential markers.) An idealization implicit[9] in all dynamic semantics frameworks is that a discourse is a chain of declarative sentences. But real discourse includes questions as well. Moreover, questions participate in many of the same dynamic processes as declaratives; although they do not assert propositions, they may introduce discourse referents, have anaphoric dependencies, force presuppositional accommodation, etc. A realistic dynamic semantics, then, will have to make room for sentences whose propositional content is not asserted, but plays some other rôle in the discourse. Question markers and evidentials like *must* are overt morphological markers of such rôles. Semantically, these morphemes are not functions from propositions to new propositions, but from propositions to labeled propositions.

The principle conclusions of this paper, then, are the following. Epistemic *must* is not a universal quantifier over possibilities, and it is not the dual of epistemic *might*, in contrast to non-epistemic *must* and *might*. It presents a proposition as being a (non-logical) deduction from some evidence that is directly known, rather than being itself directly known. This entails a rather radical re-conception of the nature of natural language semantics. The lexically-driven, context-independent semantics of a sentence includes not only propositional information about discourse referents and their properties, but meta-propositional information about the discourse rôle of propositions. A formal theory of natural language semantics has a responsibility to take this into account.

9 Hans Kamp made this assumption explicit in his presentation at the colloquium.

References

Beaver, D. 1993. *What Comes First in Dynamic Semantics: a compositional, non-representational account of natural language presupposition*. ILLC.

van Benthem, J. 1985. *A Manual of Intensional Logic*. CSLI.

Galliers, J. 1991. "Cooperative interaction as strategic belief revision". In S. Deen (Ed.). *CKBS '90: Proceedings of the International Working Conference on Cooperative Knowledge Based Systems*. Springer.

Groenendijk, J., M. Stokhof & F. Veltman. 1994. "This might be it". *ILLC Research Report and Technical Notes Series* LP-94-13. ILLC.

Heim, I. 1982. *On the Semantics of Definite and Indefinite Noun Phrases*. PhD dissertation. University of Massachusetts at Amherst.

Kartunnen, L. 1973. "Presuppositions of compound sentences". *Linguistic Inquiry* 4.

Kratzer, A. 1976. "Was ‚können' und ‚müssen' bedeuten können müssen". *Linguistische Berichte* 42.

1977. "What 'must' and 'can' must and can mean". *Linguistics and Philosophy* 1.

1981. "The notional category of modality". In H.-J. Eikmeyer & H. Rieser (Eds.). *Words, Worlds, and Contexts: New approaches in word semantics*. Walter de Gruyter.

Matthews, G. 1965. *Hidatsa Syntax*. Mouton.

Palmer, F. 1986. *Mood and Modality*. Cambridge University Press.

Veltman, F. To appear in *Linguistics and Philosophy*. "Defaults in Update Semantics".

Walker, M. 1992. "Redundancy in collaborative dialogue". *Fourteenth International Conference on Computational Linguistics*.

The Square of Individuals

Yoad Winter
OTS, Utrecht University

1 Introduction

Coordinate, plural and predicative constructions of noun phrases are common and highly productive in natural languages. They are also related in intricate ways: conjunctive coordination is one of the simple means to form plural NP's from singular ones; many coordinate and plural NP's appear naturally in predicative positions. Any theory of grammar and semantic interpretation should account for these facts. The present paper is a contribution to the semantic part of the theory. Following Partee's programmatic article *"Noun phrase interpretation and type shifting principles"* (1987) we will further explore and substantiate the paradigm of *Type Flexibility*. Extending the treatment in Winter (1994) to plural and predicative NP's, it will be shown how the long-standing puzzle surrounding "intersective" vs. "collective" conjunction can be not only resolved, but also lead to a more comprehensive theory of flexibility. The conjunctive coordinator *and* is uniformly treated as standard Boolean intersection, eliminating the common stipulation about its ambiguity. It will be proposed that one underlying property of Boolean domains, called *The Principal Filter Property* governs the apparent "non-Boolean" behaviour of conjunction. This property will lead to an extension of Partee's flexibility diagram ("The Partee Triangle") into a polymorphic scheme, *The Square of Individuals*. The advantages of this extended mechanism in treating coordination, plurality and predication will be exemplified.

2 Thoughts about non-laws of non-Boolean investigations

Modern semantics of coordination has benefited a lot from the simple and powerful Boolean view on natural language. According to this perspective, whose foundations were extensively studied in Keenan & Faltz (1985), the coordinators *and* and *or* function semantically like the operators *meet* and *join* of standard Boolean algebra. The cross-categorial behaviour of coordination is predicted in a straightforward and elegant way. To consider only one example, the Boolean definition of *and* accounts for the semantics of (1) in the propositional domain (D_t) just like its function in the domain of predicates (D_{et}) accounts for (2).

(1) Mary is tall and John is thin.

(2) Mary is tall and thin.

The Generalized Quantifier treatment of NP's following Montague allowed an equally elegant treatment of NP coordination as exemplified in (3). However, cases of "collective" interpretation like (4) seem problematic also for the Montague and GQ tradition.

(3) Mary and John smiled.

(4) Mary and John met.

Because of this last problem, most familiar treatments of plurality renounce the Boolean hypothesis. Unfortunately, this is often done without offering a plausible alternative. It has been simply claimed that cases like (4) (and possibly (3)) are evidence for "non-Boolean *and*". The common wisdom has been to assign *and* a second lexical entry that "amalgamates" singular individuals into a plural individual. This can be technically done using non-Boolean operations like i-sum or group formation, or, somewhat surprisingly, using Boolean *join*, the standard denotation of *or*.

In Winter (1994) I argue in detail against this practice. Without reviewing the more technical arguments, let us reconsider briefly two of the general objections. The first objection is methodological and (cross-)linguistic: no substantial independent motivation was given for the ambiguity of *and* in English and its parallels in numerous other languages where conjunction behaves the same. No language was shown to have two different morphemes for conjunction that function semantically in the way that is expected if *and* in English is coincidentally ambiguous. In the absence of such arguments the acclaimed ambiguity of *and* must be considered a brute force stipulation.

A second objection is on phenomenological grounds. The common wisdom is to describe the "Boolean" behaviour of *and* in (2) in terms of the equivalence in (5) and the acclaimed "non-Boolean" behaviour of *and* in (4) in terms of the *lack* of equivalence in (6).

(5) Mary is tall and thin $\Leftrightarrow$ Mary is tall and Mary is thin

(6) Mary and John met $\not\Leftrightarrow$ *?Mary met and John met

However, it is known that this "conjunction reduction equivalence" test cannot characterize two different *and*'s. One reason is that the results of the test reverse if we replace the subject in (2) by an NP with the import of an existential quantifier, and if the predicate in (4) is replaced by a "distributive" predicate (as in (3)). For example, compare (5)-(6) with (7)-(8).

(7) Some girl is tall and thin $\not\Leftrightarrow$ Some girl is tall and some girl is thin

(8) Mary and John smiled $\Leftrightarrow$ Mary smiled and John smiled

Thus, the (lack of) equivalences above is not a linguistic test distinguishing two different *and*'s. Such a general test is probably unknown.

Instead of having to face these discouraging fundamental problems, I proposed in Winter (1994) that we should better do away with the "non-Boolean" stipulation and look for a more principled account of collectivity within the *Boolean* framework. In the next section we will go through the basic Boolean property that allows the unified treatment in that earlier paper, as well as the extensions to be introduced in subsequent sections.

3 The Principal Filter Property

Intuitively speaking, in (2) conjunction seems to behave as intersection of the predicate denotations. In (4) the intuition reverses- *and* seems to function more like set *union*- taking the two singletons for *Mary* and *John* and generating one set with two elements. The "collective" predicate seems to apply to this plurality. Is there a way to resolve the confusion evoked by this seemingly double nature of conjunction? Fortunately, in standard Generalized Quantifier theory there is. Let us recall first the Boolean definition for principal filters in the quantifier domain.

(9) Let Y be any set included in a domain E. Then *the principal filter generated by* Y is the generalized quantifier $F_Y \stackrel{def}{=} \{X \subseteq E | Y \subseteq X\}$

One crucial feature of principal filters is that their *intersection* has the effect of set *union* at the level of the generators. This is stated as follows:

(10) **The principal filter property** (PFP): Let A and B be any sets. Then the *intersection* of the principal filters generated by A and B is the principal filter generated by the *union* of A and B. In formula: $F_A \cap F_B = F_{A \cup B}$.

Figure 1 exemplifies the PFP in a model where $A = \{a, c\}$, $B = \{b, c\}$, $A \cup B = \{a, b, c\}$ and $E = \{a, b, c, d\}$. The circles contained in the illustration for $F_{A \cup B}$ are exactly the ones appearing both in F_A and in F_B.

On the formal side the PFP comes as no surprise. It is just a result of the fact that the domains $P(E)$ (of predicates) and $P(P(E))$ (of generalized quantifiers) with standard set operations are Boolean algebras. In general, the PFP is a manifestation of the following Boolean algebraic property:

(11) $A \leq X$ and $B \leq X$ *iff* $A \vee B \leq X$

Consider the Boolean semantics of the propositional calculus. (11) expresses the classical equivalence $(A \rightarrow X) \wedge (B \rightarrow X) \Leftrightarrow ((A \vee B) \rightarrow X)$. In general, we can use a Chomskyan expression and say that (11) "comes for free" in Boolean semantics. Especially, we have the PFP (10) in the standard Boolean GQ semantics.

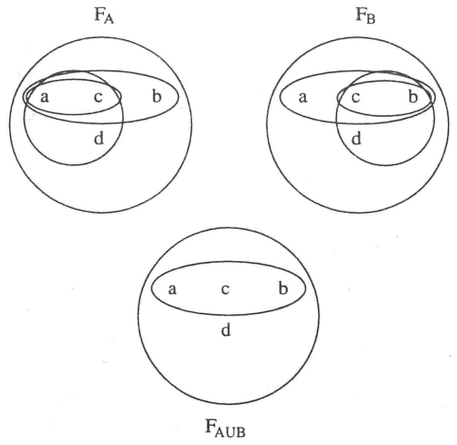


Figure 1: An example of the Principal Filter Property

On the linguistic side, the present paper continues to explore the following hypothesis about "non-Boolean" cases like (4):

(12) **Hypothesis**: "Union" behaviour of *and* in nominal conjunction is a result of the principal filter property of its standard *intersective* definition in Boolean domains.

To study better the possible ramifications and applications of the PFP hypothesis we should examine some cases of collectivity and their treatment in the proposed approach.

4 Collectivity and flexibility: some examples

On its own, the PFP is just a simple observation on standard generalized quantifiers. The problem sentences like (4) pose to the GQ tradition is not automatically resolved by it. However, let us be explicit about where the PFP is observed in the standard analysis of the subject:

$$[\![\text{Mary and John}]\!] = \{X|\{\mathbf{m}'\} \subseteq X\} \cap \{X|\{\mathbf{j}'\} \subseteq X\} \stackrel{\text{PFP}}{=} \{X|\{\mathbf{m}', \mathbf{j}'\} \subseteq X\}$$

We see that the set $\{\mathbf{m}', \mathbf{j}'\}$, which is required to get collectivity right, "hides" in the principal filter denotation of the subject: it is its generator. How can we use this fact in order to analyze correctly the collectivity in (4)? To see that, let us first review a basic assumption about plurality in Boolean frameworks (e.g. Bennett (1974), Scha (1981), van der Does (1992)). In

Bennett's approach, the domain of singular individuals is standardly D_e. This means that the domains for singular predicates and quantifiers are also standard: D_{et} and $D_{(et)t}$. Plural individuals are viewed as sets: they are (essentially) interpreted in the domain D_{et}. This causes a natural change in the types of the domains for plural predicates ($D_{(et)t}$) and plural quantifiers ($D_{((et)t)t}$). This typing strategy is summarized in table 1.

	individual	**predicate**	**quantifier**
singular	e	et	$(et)t$
plural	et	$(et)t$	$((et)t)t$

Table 1: Set theoretical plurality

The proposal to treat plural individuals as sets is minimalistic and convenient for the purpose of this paper. However, readers whose favorite ontology of plurals is different should not be too bothered. In section 6 we will see that the main results of this paper hold also with a Linkian ontology.

Getting back to (4), in the set theoretical typing the collective predicate *to meet* is given the type $(et)t$. The basic type of the singular quantifier representation for the subject is the same. In Winter (1994) I concluded that this type mismatch in functional application motivates lifting of the subject denotation into a plural quantifier. This is the crucial point at which the PFP is used to derive collectivity. It includes two steps:

1. Using the *minimum* sort operator (min) to "detect" the generator set of the principal filter. A minimum of a generalized quantifier is the set of sets in its extension that do not properly contain any other sets in the quantifier:

   $min_{((et)t)((et)t)} \stackrel{def}{=} \lambda Q_{(et)t}.\lambda P_{et}.Q(P) \wedge \forall A_{et} \subseteq P(Q(A) \rightarrow A = P)$

   It is easy to see that $min(\{X|\{\mathbf{m}', \mathbf{j}'\} \subseteq X\}) = \{\{\mathbf{m}', \mathbf{j}'\}\}$

2. Using the *existential* lifting operator (**E**) to derive a plural quantifier from the minimal sets. This operator is standardly defined by:

   $$\mathbf{E}_{((et)t)(((et)t)t)} \stackrel{def}{=} \lambda \mathcal{P}_{(et)t}.\lambda \mathcal{A}_{(et)t}.\exists P_{et}[\mathcal{P}(P) \wedge \mathcal{A}(P)]$$

   We see that $\mathbf{E}(\{\{\mathbf{m}', \mathbf{j}'\}\}) = \lambda \mathcal{A}_{(et)t}.\mathcal{A}(\{\mathbf{m}', \mathbf{j}'\})$

   At this stage predication with $\mathbf{meet}'_{(et)t}$ can obtain and we get: $\mathbf{meet}'(\{\mathbf{m}', \mathbf{j}'\})$, as required.

In the same way we can deal with more complex cases of coordination with mixed conjunction and disjunction. For example:

(13) Mary and [John or Bill] met.

$[\![\text{Mary and [John or Bill]}]\!] = \{X|\{m', j'\} \subseteq X \vee \{m', b'\} \subseteq X\}$
$\xrightarrow{min} \{\{m', j'\}, \{m', b'\}\}$
$\xrightarrow{E} \lambda\mathcal{A}_{(et)t}.\mathcal{A}(\{m', j'\}) \vee \mathcal{A}(\{m', b'\})$
Result: $\mathbf{meet}'(\{m', j'\}) \vee \mathbf{meet}'(\{m', b'\})$

Although we are not dealing here with principal filters but with a union of principal filters, the min operator can still be used to derive the generators of the filters involved. In Winter (1994) more cases of singular proper name and quantified NP coordination are analyzed in detail. However, in that paper I refrained from dealing with morphologically plural conjuncts. Let us consider now the simple case in (14).

(14) The girls and the boys met.

This sentence can be true in a situation where the girls had one meeting and the boys had a separate meeting. This kind of verifying situation is unproblematic. Let the NP's *the girls* and *the boys* be represented by the sets G'_{et} and B'_{et} respectively. Standard lifting into plural quantifiers and conjunction lead to the plausible reading $\mathbf{meet}'(G') \wedge \mathbf{meet}'(B')$.

However, there is a more problematic kind of situations that verify (14). The sentence can be true if the girls and the boys had a meeting *together*. Once again, *and* seems to behave as set union, generating the set $G' \cup B'$. According to the PFP hypothesis this effect should appear because two principal filters are conjoined. Mechanically, a natural way to guarantee that is to lift the plural individuals G' and B' into *singular* principal filter quantifiers of type $(et)t$ using the universal lifting operator. This operator is defined as follows:

$$\mathbf{A}_{(et)((et)t)} \stackrel{def}{=} \lambda P_{et}.\lambda A_{et}.\forall x[P(x) \rightarrow A(x)]$$

If we allow **A** to apply before coordinating the denotations of the conjuncts we derive two principal filters. Then the same analysis of (4) applies in (14):

$[\![\text{the girls}]\!] = G' \xrightarrow{\mathbf{A}} \{X|G' \subseteq X\}$
$[\![\text{the boys}]\!] = B' \xrightarrow{\mathbf{A}} \{X|B' \subseteq X\}$
$[\![\text{the girls and the boys}]\!] = \{X|G' \subseteq X\} \cap \{X|B' \subseteq X\} \stackrel{PFP}{=}$
$\{X|G' \cup B' \subseteq X\}$
$\xrightarrow{min} \{G' \cup B'\}$
$\xrightarrow{E} \lambda\mathcal{A}_{(et)t}.\mathcal{A}(G' \cup B')$
Result: $\mathbf{meet}'(G' \cup B')$, as required.

This application of **A** allows a correct extension of many results from Winter (1994) to the domain of plurality. However, instead of doing that

too hastily, I think there is a more fundamental point to deal with first. The devil's advocate must be heard at this stage.

One main criticism against the analysis sketched above is along the following lines. In search of a unified treatment of conjunction one stipulation ("non-Boolean" *and*) seems to be replaced by three (flexibility operators). What do we gain by this move? Answering this completely well-founded objection is the main aim of this paper. It will be shown that if we follow Partee (1987), the three operators min, **E** and **A** can be viewed as the core engine of a general flexibility mechanism, replacing the three main operators in Partee's treatment. Partee's operators were motivated on independent grounds for treating singular NP's. This will imply that also the three operators used above in the analysis of plurality are independently motivated. Some of these motivations will be additional to the ones discussed by Partee. The proposed revision of Partee's diagram will therefore be shown to explain more facts using a similar amount of assumptions. Thus, it will be proposed that a unified account of conjunction as reviewed above can lead towards *a theory of flexibility*, accounting also for the interpretation of plural and predicative NP's.

5 Squaring the Partee Triangle

Partee (1987) argues for a significant change in the classical Montagovian view on the syntax-semantics interface. Instead of Montague's "strict" assignment of a unique semantic typed denotation to each syntactic category, Partee proposes a more "flexible" strategy. Following Partee & Rooth (1983), NP's are allowed to be interpreted using more than one denotation. From unambiguous lexical denotations multiple interpretations of expressions can be derived using available flexibility operators that apply in the process of the compositional interpretation.

One of the major contributions of Partee's paper is the emphasis it puts on the following theoretical point. Flexible semantics should not be just an unconstrained arsenal of polymorphisms. We would like the mechanism to be empirically motivated but also theoretically restrictive. There should be some *order* in the array of flexibility operators proposed. The order revealed by Partee was described using a simple diagram, sometimes referred to in the literature as *The Partee Triangle* (PTR). This diagram and its relevant linguistic applications are now to be briefly reviewed.

5.1 The Partee Triangle

In figure 2 appears a simplified version of the PTR. Some operators which will not be relevant for our discussion are omitted. The PTR includes the three main types proposed for singular NP's: e (individuals), et (predicates) and $(et)t$ (quantifiers). Three main functions connect denotations

in these types. An individual can be lifted to the corresponding singleton predicate using the operator **Q** (Quine). A predicate can be lifted to the corresponding existential quantifier using **E** (Existential). A quantifier can be lowered to a predicate using Montague's definition **B** (Be) for the copula in English. A composition of **Q** and **E** shifts an e type individual into the corresponding quantificational principal ultrafilter. We abbreviate $\mathbf{E} \circ \mathbf{Q}$ by **M** (Montague Raising). For symmetry (although not necessarily for linguistic reasons) the PTR includes also the inverse of **M**, the partial function $\mathbf{M}^{-1}$ (Montague Lowering). The exact definitions are given below. Some interesting relations between these operators were studied in Partee (1987) and van Benthem (1986).

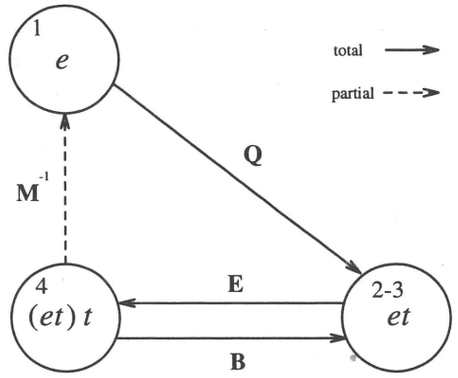


Figure 2: The Partee Triangle

Quine:	$\mathbf{Q}_{e(et)}$	$\stackrel{def}{=}$	$\lambda x_e.\lambda y_e.x = y$
Existential:	$\mathbf{E}_{(et)((et)t)}$	$\stackrel{def}{=}$	$\lambda P_{et}.\lambda A_{et}.\exists x_e(P(x) \wedge A(x))$
Be:	$\mathbf{B}_{((et)t)(et)}$	$\stackrel{def}{=}$	$\lambda Q_{(et)t}.\lambda x_e.Q(\{x\})$
M. Raising:	$\mathbf{M}_{e((et)t)}$	$=$	$\mathbf{E} \circ \mathbf{Q} = \lambda x_e.\lambda P_{et}.P(x)$
M. Lowering:	$\mathbf{M}^{-1}_{((et)t)e}$	$\stackrel{def}{=}$	$\lambda Q_{(et)t}.\begin{cases} a_e & Q = \mathbf{M}(a) \\ \text{undefined} & \text{otherwise} \end{cases}$

The Quine operator is useful for describing the behaviour of individual denoting NP's when appearing in predicative positions. For example, in (15) **Q** derives the appropriate predicate.

(15) This teacher is John.

$\mathbf{j}'_e \xrightarrow{\mathbf{Q}} \lambda x.x = \mathbf{j}'$

Result: $[\![\text{this teacher}]\!]_e = \mathbf{j}'$

The Existential operator derives quantificational readings of indefinites. This can be done in classical Montague semantics using the extensional denotation of the indefinite article. However, in many languages indefinites can appear without any overt article. An insight that Partee shares with DRT and other semantic frameworks is the idea that "existential closure" is a semantic operation which does not necessarily correspond to a lexical denotation. It is required in order to get application right when indefinites appear in argument positions, and this is because indefinites basically denote semantic *predicates*. For example, the Hebrew sentence (16) can be analyzed using **E**, although no morpheme with this determiner denotation overtly appears in the sentence. The analysis of the English counterpart is parallel, using the assumption that the overt article in English lacks "ordinary" denotational contribution to meaning.

(16) adam nixnas laxeder

person entered into-the-room

"A person entered the room"

$\mathbf{person}'_{et} \xrightarrow{\mathbf{E}} \lambda P_{et}.\exists x_e(P(x) \wedge \mathbf{person}'(x))$

Result: $\exists x[\mathbf{enter}'_{et}(x) \wedge \mathbf{person}'(x)]$

The motivation for the **B** operator is to allow predicative readings of quantifiers. For example, once the predicate in (17) is lifted to an existential quantifier it can be lowered back to get the same predicate.

(17) John is a teacher.

$\mathbf{teacher}'_{et} \xrightarrow{\mathbf{E}} \lambda P_{et}.\exists x_e(P(x) \wedge \mathbf{teacher}'(x)) \xrightarrow{\mathbf{B}} \mathbf{teacher}'$

Result: $\mathbf{teacher}'(\mathbf{j}')$

This seems somewhat redundant. If indefinites are basically treated as predicates the same result in (17) can be obtained directly. This familiar position to predication with indefinites will be adopted below. However, a more interesting case is the following:

(18) John is no teacher.

$[\![\text{no teacher}]\!]_{(et)t} = \lambda P.P \cap \mathbf{teacher}' = \emptyset \xrightarrow{\mathbf{B}} \lambda x.\neg\mathbf{teacher}'(x)$

Result: $\neg\mathbf{teacher}'(\mathbf{j}')$

The well-known application of **M** ($= \mathbf{E} \circ \mathbf{Q}$) is to allow the analysis of coordination with proper names. In (extensional) Montague grammar individuals are lexically given the conjoinable type $(et)t$. In the PTR, the required lifting can be achieved also operationally while singular individuals are assigned the type e.

5.2 The Square of Individuals

We are in a position to introduce the main intuition underlying the proposed extension of the Partee Triangle. The three operators **A**, **E** and min that were used in section 4 will be conceived as follows:

- The universal lifting operator **A** is the available mapping from individuals to quantifiers.
- The existential lifting operator **E** is the available mapping from predicates to quantifiers.
- The minimum sort operator min is the available mapping from quantifiers to predicates.

Note that in this conception, as necessitated by the Bennett typing strategy, types alone do not tell us exactly whether the denotation is of an individual, a predicate or a quantifier. For example, et is the type of both plural individuals and singular predicates. It is therefore more instructive to think of individuals being basically of an arbitrary type τ (not necessarily e), where the corresponding predicates and quantifiers over this type of individuals get the types τt and $(\tau t)t$.

This conception underlies the configuration of *The Square of Individuals* (SQI). The mechanism is illustrated in figure 3 and the participating operators are defined below. A polymorphic definition for the inclusion relation $\sqsubseteq_{\tau(\tau\tau)}$ is assumed and will be given in what follows (cf. (23)).

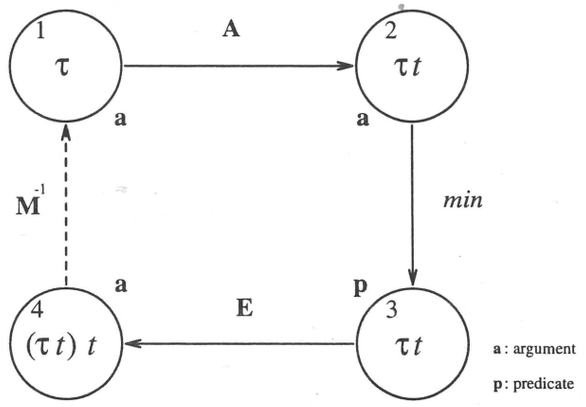


Figure 3: The Square of Individuals

Universal:	$\mathbf{A}_{\tau(\tau t)}$	$\stackrel{def}{=}$	$\lambda x_\tau.\lambda y_\tau.x \sqsubseteq y$
Minimum:	$min_{(\tau t)(\tau t)}$	$\stackrel{def}{=}$	$\lambda P_{\tau t}.\lambda x_\tau.P(x) \wedge \forall y \sqsubseteq x(P(y) \rightarrow y = x)$
Existential:	$\mathbf{E}_{(\tau t)((\tau t)t)}$	$\stackrel{def}{=}$	$\lambda P_{\tau t}.\lambda A_{\tau t}.\exists x_\tau(P(x) \wedge A(x))$
M. R.:	$\mathbf{M}_{\tau((\tau t)t)}$	$=$	$\mathbf{E} \circ min \circ \mathbf{A} = \lambda x_\tau.\lambda P_{\tau t}.P(x)$
M. L.:	$\mathbf{M}^{-1}_{((\tau t)t)\tau}$	$\stackrel{def}{=}$	$\lambda Q_{(\tau t)t}.\begin{cases} a_\tau & Q = \mathbf{M}(a) \\ \text{undefined} & \text{otherwise} \end{cases}$

Before getting into some substantial matters concerning the SQI, let us note two independent changes made in the PTR in order to deal with plurality and predication. One change is the above mentioned parameterization of the type for individuals. This is conceptually convenient when making the type distinction between singular (e type) individuals and plural (et type). These are the only types which will be used for τ. The parameterization is made only for expository reasons. Implementations of the SQI that will be considered in section 6 instantiate the type for τ.

A second general addition to the PTR is needed in order to be more explicit about the function of the denotation derived by the predicative vertex of the scheme. In the SQI a separate feature **a/p** of a vertex signals if the NP denotation derived is reserved to NP's in syntactic argument/predicate positions. In the PTR this information does not appear, which creates some problems in the analysis of coordination. For example, in order to analyze the simple sentence (19) correctly, we should block the derivation (19a), where the coordinated denotations are predicative (type et).

(19) Mary and John smiled.

a. * $\mathbf{m}'_e \xrightarrow{\mathbf{Q}} \{\mathbf{m}'\}$ $\mathbf{j}'_e \xrightarrow{\mathbf{Q}} \{\mathbf{j}'\}$
$\{\mathbf{m}'\} \cap \{\mathbf{j}'\} = \emptyset_{et}$
$\emptyset_{et} \xrightarrow{\mathbf{E}} \emptyset_{(et)t}$
$\emptyset_{(et)t}(\mathbf{smile}'_{et}) = \mathit{false}$

This contradictory result is of course undesirable. I follow Partee in the intuition that the predicative reading of NP's is "marked" and reserved for NP's in syntactic predicative positions (see Williams (1983)). This means that in an argument position, the NP is neither interpreted nor coordinated if its denotation was last derived in a vertex marked with the feature **p**. With the appropriate mechanism this blocks the derivation in (19a). This move is unattractive, but I do not know of any general account of predication that manages to avoid it. Thus, at the moment we can consider the step necessary, irrespective of the flexible mechanism proposed.

Let us consider now the SQI itself. I think four main points can fairly summarize the new aspects of the proposed analysis:

(i) The flexibility operator **A** is motivated as a *distributivity* operator proposed in previous works on plurality.

(ii) The same operator is used here to derive *collectivity* effects in coordinations, given the PFP hypothesis.

(iii) The minimum operator allows NP's to have a *collective predicative* denotation.

(iv) Montague's **B** operator is eliminated.

Some elaborations follow.

"Distributivity as flexibility" Probably all contemporary accounts of plurality use some special mechanism to derive distributivity effects. The reason is that plural NP's often behave as *individuals* being in the extension of a "collective" predicate as in (20). But they can also have a *quantificational* ("distributive") import as in (21), intuitively equivalent to (21a).

(20) The girls met.

(21) The girls are wearing a nice dress.

a. *Every* girl is wearing a nice dress.

b. $[\![\text{the girls}]\!] = G'_{et} \xrightarrow{\mathbf{A}} \lambda P_{et}.\forall x[G'(x) \rightarrow P(x)]$
Result: $\forall x[G'(x) \rightarrow \exists y(\mathbf{dress}'(y) \wedge \mathbf{wear}'(y)(x))]$

A distributivity operator is commonly introduced into the system to allow the derivation of the quantificational effect. We can ignore here the popular debate whether a distributivity operator should be located on the NP argument or on the VP predicate. In (21b) and in what follows it is assumed that the first strategy is at least a possible derivation. The same position was taken in Heim, Lasnik & May (1991), among others. The new proposal here is that distributivity is a *flexibility* operation that should be part of a general scheme like the PTR. The SQI incorporates this view, one surprising consequence of which can now be noticed.

"Distributivity for collectivity" Recall that the **A** operator was required in the analysis of (14) because of the PFP hypothesis. In order to get the "union"/"collective" function of Boolean *and* we had to obtain principal filters as conjuncts. **A**, a traditional operator of distributivity is responsible also for this "collective" effect. This is a consequence of viewing **A** as a general operator for flexible NP interpretation. It seems to add important support for its use, which is commonly only motivated by the mechanical need to derive distributivity for plural individuals.

"Collectivity in predication" In the PTR the operators **Q** and **B** are responsible for predicative interpretation of NP's. In the SQI the minimum operator is the only one that reaches the predicative vertex. For example, in the singular case where $\tau = e$, the composition $min \circ \mathbf{A}$ is equivalent to the **Q** operator of the PTR (see below). This allows the correct treatment of sentences like (15) also in the SQI. When we deal with proper name conjunction using standard Boolean *and*, some differences between the PTR and the SQI appear. Consider a conjunction like *John and Bill* in a predicative position as in (22). In the PTR there are two possibilities to analyze the sentence: either with **B** lowering of the quantificational reading of the conjunction or by conjoining the predicative denotations obtained by **Q**. Both possibilities, given in (22a) and (22b), lead to an empty predicate that makes (22) be analyzed as contradictory.

(22) These men are John and Bill.

a. $\mathbf{j}' \xrightarrow{\mathbf{M}} \lambda P_{et}.P(\mathbf{j}') \quad \mathbf{b}' \xrightarrow{\mathbf{M}} \lambda P_{et}.P(\mathbf{b}')$
$\lambda P.P(\mathbf{j}') \wedge P(\mathbf{b}') \xrightarrow{\mathbf{B}} \emptyset_{et}$

b. $\mathbf{j}' \xrightarrow{\mathbf{Q}} \{\mathbf{j}'\} \quad \mathbf{b}' \xrightarrow{\mathbf{Q}} \{\mathbf{b}'\}$
$\{\mathbf{j}'\} \cap \{\mathbf{b}'\} = \emptyset_{et}$

These are not necessarily bad results. Partee claims that such derivations might reflect a distributive reading of the predicative NP, which is here necessarily false: the group referred to by the subject cannot be *both* John and Bill. However, (22) has a *contingent* reading which is not captured by the PTR without further assumptions. In the SQI this reading is derived:

$\mathbf{j}' \xrightarrow{\mathbf{M}} \lambda P_{et}.P(\mathbf{j}') \qquad \mathbf{b}' \xrightarrow{\mathbf{M}} \lambda P_{et}.P(\mathbf{b}')$
$\lambda P.P(\mathbf{j}') \wedge P(\mathbf{b}') \xrightarrow{min} \lambda P_{et}.P = \{\mathbf{j}', \mathbf{b}'\}$
Result: $[\![\text{these men}]\!]_{et} = \{\mathbf{j}', \mathbf{b}'\}$

This shows a point where the general treatment of predicative NP's in PTR (and classical Montague grammar) is revised. The change can be described using the following pessimistic (although in the end admittably true) resolution to the Shakespearean dilemma.

"Not to B!" In the SQI, as its name implies, we can deal with predicative readings of (singular and plural) individual denoting NP's and their coordinations, but *not* with arbitrary generalized quantifiers. This point should be emphasized and it will be further discussed below. As we noted, sentences like (17) with a predicative indefinite NP do not require any special predicative mechanism (also in the PTR) because the basic denotation of indefinites is predicative. If this view is correct, Montague's **B** operator loses much of its motivation. But now cases like (18), (which Partee considers somehow

marked in English) become problematic in the SQI. I have no solution of this problem to offer at this stage. However, the hypothesis that predicative NP's are only NP's that *basically* denote an individual or a predicate seems to be supported by the work in Williams (1983) and Higginbotham (1987). Only a better study of this hypothesis can show what should be the treatment of such problematic cases. Of course, a last resort can always be adding **B** to the SQI. But such a theoretically expensive move should in any case have independent support. This was not given to my knowledge also for Montague's original treatment of predication.

The discussion above concentrates on some of the main differences between the PTR and the SQI. However, there is also a central point of similarity: once the operator **B** is eliminated it can be observed that the SQI formally *generalizes* the PTR. This is transparent when using van Benthem (1986)'s polymorphic definition for inclusion:

$$(23)\quad X \sqsubseteq Y \stackrel{def}{=} \begin{cases} X = Y & \tau = e \\ X \rightarrow Y & \tau = t \\ \forall Z_{\sigma_1}(X(Z) \sqsubseteq Y(Z)) & \tau = \sigma_1\sigma_2 \end{cases}$$

The intuition behind this definition is that inclusion gets the standard set theoretical definition in Boolean (t-reducible) domains. In non-Boolean (e-reducible) domains we use the only general relation available between entities– identity. This preserves "as much Boolean structure as possible" in non-Boolean domains. Using definition (23) we get the following results when substituting $\tau = e$ in the SQI:

- $min_{(et)(et)} = \lambda P_{et}.\lambda x_e.P(x)$ (identity for predicates)
- $\mathbf{A}_{e(et)} = \mathbf{Q}$

Thus, vertices 2 and 3 in the SQI (figure 3) are mapped to vertex 2-3 in the PTR (figure 2).

Let us conclude. On the formal side, the SQI lifts individuals to quantifiers using **A**, instead of mapping them to predicates using **Q** as in the PTR. The operator min of the SQI sorts quantifiers created from individuals and their coordinations to predicates, instead of the general type lowering **B** of the PTR. Linguistically, these formal distinctions allow to derive quantificational distributivity using **A**, predicative collectivity using min and a unified Boolean treatment of coordination using the combination of the operators, as expected by the hypothesis about the Principal Filter Property. These points are summarized in table 2 below.

6 Implementation: Triangulating the Square of Individuals

Conceptually, the SQI is a convenient way to present the proposed ideas about flexibility. However, two problems of implementation need to be

			PTR	SQI
individual	→	predicate	$\mathbf{Q}$ [1]	–
individual	→	quantifier	–	$\mathbf{A}$ [2,3]
predicate	→	quantifier	$\mathbf{E}$ [4]	$\mathbf{E}$ [5,6]
quantifier	→	predicate	$\mathbf{B}$ [1]	min [1,6]
individual	→	quantifier	$\mathbf{E} \circ \mathbf{Q} = \mathbf{M}$ [7]	$\mathbf{E} \circ min \circ \mathbf{A} = \mathbf{M}$ [7]
quantifier	→	individual	$\mathbf{M}^{-1}$ [8]	$\mathbf{M}^{-1}$ [9]

[1] predicative NP's
[2] distributivity
[3] PFP (deriving collectivity)
[4] indefinites
[5] indefinites (but see below)
[6] deriving collectivity
[7] proper names in coordination
[8] not required?
[9] not required (see below)

Table 2: Differences between the PTR and SQI and applications

treated. The first problem has to do with the type parameterized domain of individuals. In a standard type theoretical implementation of the SQI we have to instantiate τ according to the ontology in our semantics. I will propose two possible directions. (**i**) Using a set theoretical approach to plurality, we will instantiate τ as e and et and merge the two resulting diagrams. Evidently, this move is much in accordance with the ontology minimalistic line adopted in this paper. We will follow Keenan & Faltz (1985) (and ultimately Montague) in eliminating reference to the domain of individuals D_τ from the semantic mechanism. (**ii**) Enriching the ontology of individuals with pluralities related to their parts using some non-Boolean structuring of the e domain (a Linkian lattice treatment, for example).

In both approaches we solve the following problem, where the SQI as presented above is fundamentally inadequate. Given the assumption that plural individuals are treated as et type entities, we expect them to be coordinated using the Boolean coordinators. This is evidently incorrect. To consider one simple example, we expect (24) to be interpreted as (24a), which is clearly not the case.

(24) The girls *or* the boys met.
   a. * $\mathbf{meet}'(G' \cup B')$

In a set theoretical implementation, individuals should be directly interpreted in the $(\tau t)t$ type, so this problem is avoided. In a lattice theoretical approach, all individuals are e type, so the problem does not appear in the first place. Let us elaborate on these points.

6.1 Set (Scha-Bennett) theoretical plurality

As said above, in the S-theoretical approach we follow Keenan & Faltz (1985) and "eliminate the universe" from the SQI. All individuals are directly interpreted in the $(\tau t)t$ domain. Since **A** should apply now to principal ultrafilters instead of plain individuals it is composed with $\mathbf{M}^{-1}$ of the SQI to get the triangle in figure 4. Now we can safely instantiate τ as e and et. To do that, we use the standard definition of Boolean inclusion in (25) (not van Benthem's generalization (23) with the double base for the recursion). In this definition **A** and min of the SQI are undefined when $\tau = e$. [1] Consequently, in the e type instantiation vertex 2 of the SQI becomes unreachable and it is therefore eliminated from the scheme.

(25) $X \sqsubseteq Y \stackrel{def}{=} \begin{cases} X \to Y & \tau = t \\ \forall Z_{\sigma_1}(X(Z) \sqsubseteq Y(Z)) & \tau = \sigma_1\sigma_2 \end{cases}$

After type instantiation, the squared-triangulated-squared-Partee-triangle in figure 4 is derived. Thus, we end up with triangles and squares even.

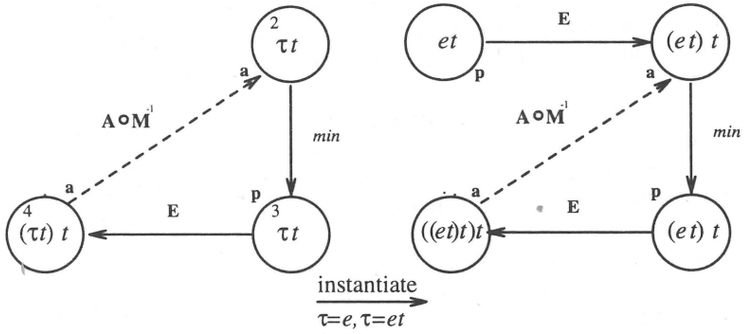


Figure 4: An S-plurality instantiation

A consequence of this implementation is that predicative NP's cannot be of type et, but only of type $(et)t$. Thus, simple sentences like (15) must be analyzed using a lifting of the subject, as exemplified in (26). Albeit somewhat unintuitive, this seems unproblematic.

[1] One may wonder what is the justification for not using here van Benthem's definition while using it above in showing the relations between the PTR and the SQI. The way I see it, (23) is convenient for meta-theoretical uses, when wanting to check the relations between Boolean and non-Boolean domains. However, for practical uses, it was noted in Winter (1994) that (23) is expected to derive definitions for conjunction and disjunction in non-Boolean domains which are linguistically unintuitive. Therefore, while dealing with actual implementation the classical definition (25) is preferable.

(26) This man is John.

$[\![\text{this man}]\!]_{(et)t} = \lambda P_{et}.P(a_e) \stackrel{\text{E}\circ min}{\longrightarrow} \lambda\mathcal{P}_{(et)t}.\mathcal{P}(\{a\})$

$[\![\text{John}]\!] = \lambda P_{et}.P(\mathbf{j}') \stackrel{min}{\longrightarrow} \lambda A_{et}.A = \{\mathbf{j}'\}$

Result: $\{a\} = \{\mathbf{j}'\} \Leftrightarrow a = \mathbf{j}'$

6.2 Lattice (Link) theoretical plurality

In the L-plurality approach there are no et individuals, but there are some assumptions about the structure of the e domain. Following Partee (1987), we can view L-ontology as assuming a primitive mapping from type e to et, a mapping that Partee calls *delink*:

(27) $delink_{e(et)}(x) = S$ iff S is the collection of atomic i-parts of x

In order to implement the SQI in this approach we can use *delink* by revising the definitions of **A** and min as follows:

$\mathbf{A}^L \stackrel{def}{=} \lambda x_e.\lambda P_{et}.\mathbf{A}(delink(x))(P)$

$min^L \stackrel{def}{=} \lambda Q_{(et)t}.\lambda x_e.min(Q)(delink(x))$

The SQI is triangulated now as in figure 5.

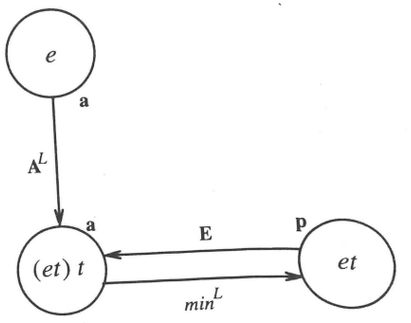


Figure 5: An L-plurality implementation

A possible drawback is that **M** is no longer generally derivable:

$\mathbf{M}(x) = \mathbf{A}^L(x)$ *only if* x is atomic ($|delink(x)| = 1$)

Thus, raising of *plural* individuals (like the one denoting *the boys*) is not accounted for. To solve this problem, we can try to follow most versions of categorial grammar, where **M** is derived by more general categorial

principles. Another alternative is to eliminate the universe as in the S-plurality approach, to shrink the SQI into two vertices. As in the S-plurality implementation above, this configuration requires no lifting of individuals.

Instead of diving deeper into these technical swans, let us conclude with some more general problems and prospectives.

7 Choice functions and the hyper-flexibility of indefinites

In the presentation of the SQI we saw how the Principal Filter Property hypothesis independently motivates the distributivity operator **A** for plural NP's. Let us turn now briefly to another prediction about the interpretation of *indefinite* NP's. In Winter (1994,1995) it was noticed that standard treatments of indefinites as generalized quantifiers do not combine with the PFP to derive the correct interpretation of sentences like (28)-(29). The reason is that existential quantifiers are not principal filters and yet these sentences have a "collective" interpretation for the conjunction.

(28) Some author and some teacher met.

(29) Three authors and two teachers met.

In Winter (1995) I proposed that an analysis of indefinites using *choice functions* is a correct way to overcome such difficulties. This treatment, based on the logical intuition in Hilbert & Bernays (1939), was motivated in Reinhart (1992,1995). Reinhart proposes that it accounts for the peculiar scope behaviour of indefinites. The same mechanism accounts for the interpretation of (28)-(29) in the present treatment of coordination. An elaboration on some linguistic and formal perspectives can be found in Winter (1995). Without getting into technical details, let me just briefly say how this treatment allows a unified approach to the *flexibility* of indefinites.

A prominent view on indefinites regards them as *individuals*. This is the intuition in Fodor & Sag (1982), motivated by the wild wide scope behaviour of indefinites. This is probably also a part of the intuitive background for DRT (see some remarks in Partee (1987)). The well-known origin for this intuition is the properties of indefinites in discourse anaphora. In (28)-(29) we see a different motivation coming from the PFP hypothesis. Another plausible approach can be found in the intuitive discussion in Williams (1983) and Higginbotham (1987). According to these works indefinites are basically *predicates*. One motivation is the natural appearance of indefinites in predicative constructions. Yet, a straightforward evidence in favour of the Russellian/Montagovian *quantificational* approach is that indefinites can be interpreted as having narrower scope than other quantificational elements.

This probably covers the three most intuitive ways to think about indefinites compositionally. Are we confronted with a paradox? In the choice function approach the answer is negative. The reason, roughly, is that indefinites are basically treated as predicates. However, the choice mechanism can

"pick" an individual (principal filter) from the predicate. This mechanism itself is quantificational, allowed to have any scope in the compositional interpretation process of the sentence. For technical details see Winter (1995).

The variable binding mechanism in the choice function approach is not a straightforward modeltheoretic flexibility operation like existential lifting in the quantificational line. However, as shown in Winter (1995), when choice functions are interpreted compositionally the mechanism formally *extends* standard existential quantification. One point where this fact might allow us to achieve a broader perspective on flexibility will be discussed in the following section, together with another open question.

8 Two open questions

Let us conclude with two points that motivate further empirical study in order to generalize the proposals made in this paper.

One problem implied above has to do with the generality of the choice function mechanism. Since the **E** operator of the PTR and SQI does not seem general enough in order to treat indefinites correctly, it is reasonable to expect the more general mechanism of choice functions to replace it in other cases of flexibility as well. Consider the usage of **E** in the analysis of (4) and (13). To use here a choice function treatment as in Winter (1995) is technically straightforward. However, this also has some empirical implications. An important one is that the "scope" of the disjunction in cases like (13) behaves then like the scope of simple indefinites- violating islands. For example, (30) below should have the interpretation in (30a).

(30) If Mary and [John or Bill] meet then a terrible fight will take place.

a. $[\mathbf{meet}'(\{\mathbf{m}', \mathbf{j}'\}) \rightarrow \mathbf{fight}'_t] \vee [\mathbf{meet}'(\{\mathbf{m}', \mathbf{b}'\}) \rightarrow \mathbf{fight}']$

"Either John or Bill meeting Mary will cause a terrible fight"

(30a) is not an obvious reading of (30). Therefore, it seems somewhat too early to integrate the choice function mechanism into the SQI. More empirical tests should be devised in order to evaluate this move.

Another open point mentioned earlier is that the SQI should treat only individual denoting NP's and their coordinations, but not arbitrary generalized quantifiers. One simple case that exemplifies that is that we do not want NP's like *every man* to get the same treatment as *John and Bill* because the latter can be "collectivized" whereas the second cannot (cf. *?*every man met*). However, extensionally these two NP's can both refer to the same principal filter in a model where John and Bill are the only men. This points out that some syntactic information about NP's might have to be used in the general treatment of flexibility. Recent works (e.g. the papers in Szabolcsi (forthcoming)) suggest to use some syntactic differentiation of NP's in the semantic analysis. The proper way to do that within the proposed mechanism of flexibility is a major open problem.

9 Conclusions

Montague's program, based on the simple and powerful typed lambda calculus, has proved extremely fruitful in the semantic analysis of natural language. Partee's work revealed some limitations in the "strict" typing of Montague grammar. The flexibility revision Partee proposed allows to achieve better descriptive adequacy as well as conceptual elegance within the highly restrictive Montagovian framework. In this paper I tried to illustrate this point by considering a wider range of data and a more general formalism. This is necessary in order to continue developing Montague semantics into an integral part of linguistic theory.

References

Bennett, M.R.: 1974, *Some Extensions of a Montague Fragment of English*, Ph.D. diss., UCLA.

Benthem, J. v.: 1986, *Essays in Logical Semantics*, D. Reidel, Dordrecht.

Does, J. v.d.: 1992, *Applied Quantifier Logics*, Ph.D. diss., University of Amsterdam.

Fodor, J.D., and I. Sag: 1982, "Referential and quantificational indefinites", *Linguistics and Philosophy* 5, 355-398.

Heim, I., H. Lasnik and R. May: 1991, "Reciprocity and plurality", *Linguistic Inquiry* 22, 63-101.

Higginbotham, J.: 1987, "Indefiniteness and predication", in E.J. Reuland and A.G.B. ter Meulen(eds.), *The Representation of (In)definiteness*, MIT Press.

Hilbert, D., and P. Bernays: 1939, *Die Grundlagen der Mathematik II*, second edition, reprint 1970, Springer, Berlin.

Keenan, E., and L. Faltz: 1985, *Boolean Semantics for Natural Language*, Reidel, Dordrecht.

Partee, B.: 1987, "Noun phrase interpretation and type shifting principles", in J. Groenendijk et al.(eds.), *Studies in Discourse Representation Theories and the Theory of Generalized Quantifiers*, Foris, Dordrecht.

Partee, B., and M. Rooth: 1983, "Generalized conjunction and type ambiguity", in R. Bauerle et al.(eds.), *Meaning, Use and Interpretation of Language*, De Gruyter, Berlin.

Reinhart, T.: 1992, "Wh-in-situ: an apparent paradox", in P. Dekker and M. Stokhof(eds.), *Proceedings of the Eighth Amsterdam Colloquium.*

Reinhart, T.: 1995, "Quantifier scope", unpublished ms., OTS, Utrecht University. To appear in T. Reinhart, *Interface Strategies*, MIT Press.

Scha, R.: 1981, "Distributive, collective and cumulative quantification", in J. Groenendijk et al.(eds.), *Formal Methods in the Study of Language*, Amsterdam.

Szabolcsi, A.(ed.): forthcoming, *Ways of Scope Taking*, Kluwer.

Williams, E.: 1983, "Semantic vs. syntactic categories", *Linguistics and Philosophy* 6, 423-446.

Winter, Y.: 1994, "A unified semantic treatment of singular NP coordination", unpublished ms., OTS, Utrecht University. To appear in *Linguistics and Philosophy*.

Winter, Y.: 1995, "Choice functions and the scopal semantics of indefinites", unpublished ms., OTS, Utrecht University. (Revised and extended version of "On the formalization of choice functions as representing the scope of indefinites", in G.V. Morrill and R.T. Oehrle(eds.), Formal Grammar, proceedings of the conference of the European summer school in Logic, Language and Information, Barcelona, 1995). Available from the author.

A Neoclassical Analysis of Belief Sentences

Henk Zeevat
University of Amsterdam

The classical logic of belief is (Hintikka 1971). Hintikka provides a belief operator defined over possible worlds and treats quantification by quantifying over a set of individual concepts. This paper adds some modern features.

Many problems are known about possible world based belief theories and the aim of this paper is to repair a number of them. The aim is not so much to give a logic of belief, but rather an analysis of belief reports. What we are interested in, is a theory of the interpretation of belief sentences which connects to a logic containing a belief operator. The theory of interpretation makes the connection of the surface sentence to its logical representation an indirect one, determined by various contextual factors. Some of the problems that arise within the logical system in which the belief report is represented do not arise for the natural language belief reports. So what we are studying is not the logic of belief but the intuitive consequence relation between belief reports. I believe bringing in the recent advances in formal pragmatics to the study of these problems brings about a clarification of some issues around belief sentences, notably about the status of names and indexicals in opaque contexts.

What this paper adds to Hintikka's theory is first of all a new interpretation of the belief alternatives assumed by Hintikka. This analysis is Haas-Spohn's diagonal theory of belief alternatives, based on a radical indexicalism towards the meanings of natural language lexemes. Without this underpinning, the current paper could not have been written.

The second element that is added to Hintikka's theory is a different account of the individual concepts which are the basis of his account of quantifying into belief contexts. This account is in essence my earlier history-based theory of objects of belief (the most recent version is Zeevat 1994), but this time formulated in possible worlds semantics rather than in discourse representation theory . This formalisation is useful and adds to the original account.

Third, we offer an alternative account of proper indexicals. Here, we use the theory of presupposition that has been developed over the last ten years primarily by Van der Sandt (Van der Sandt 1992).

The resulting theory also offers a radically different account of existential *de dicto* belief reports. This account has an important consequence: it opens the door to a solution of the problem of closure under logical consequence, a solution that however relies on the idea that all judgements are essentially existential. This view can be found in certain versions of Davidson's approach to event semantics, in certain forms of DRT and of course in the intuitionistic tradition, especially in type theory.

The organisation of the paper is as follows. In the first section, we will introduce Haas-Spohn's theory of subjective interpretation and the consequences of this the-

ory for the analysis of belief sentences. Section 2 removes the basic obstacle to the extension of this theory to all *de dicto* belief sentences: Kaplan's theory of demonstratives. Section 3 is concerned with individual concepts and will enable us to give a full account of belief sentences. Section 4 gives a closer look at two traditional problems with belief sentences and presents a theory of names. Section 5 finally offers the tentative solution to closure under logical consequence and explores some consequences of the present theory for the theory of information and for natural language semantics.

The theory of belief that defines the truth of belief reports in terms of a set of belief alternatives for the subject suffers from a large number of problems. These problems originate from on the one hand the conception of the relation of belief and necessity and on the other from the fact that a set of possible worlds is always closed under logical consequence. Representative of the first problem is Kripke's puzzle (Kripke 1979).

Logical consequence is a problem of the connection of belief with practical reasoning. As beliefs are the basis for action, the theory commits us to actions we are not in fact prepared to carry out. Kamp (1994) gives a simple example involving an arithmetical truth.

1 A Diagonal Theory of Belief Alternatives

What information does a subject get from a conversation? Which belief alternatives can be eliminated?

Suppose you are availing yourself of the Blind Date Telephone Service and you are speaking to a girl who claims to be a blonde. The subjective information you are deriving from her claim (which you happen to believe) is not at all the same as the proposition expressed by the sentence "I am a blonde" which she whispered in the phone. That proposition is about that girl at the other side of the phone, Anita, a person you do not happen to know. The information you have cuts down your set of belief alternatives (the other hair colours) but the criterion by which you cut them down is not related to the girl herself, but to whoever happens to be on the phone according to the particular alternative. There are some limitations: she has that kind of voice, she is that sort of person, but within those limits anything goes. If the person on the phone in the alternative is a brunette, the alternative is rejected, if she is a blonde it stays. What is going on with Anita in the alternative is irrelevant.

The criterion for eliminating a belief alternative uses not the actual referent of "I", but the referent of "I" according to the belief alternative. If the "I" of the belief alternative is not a blonde, the alternative is elminated. That means, we have to diagonalise the sentence's character (in terms of Kaplan's character semantics) in order to find out which belief alternatives survive the update. The subjective meaning of the sentence for a person a can be defined as the intersection of the diagonal of the sentence's character with the belief alternatives of a. This analysis of subjective information is due to Haas-Spohn (following ideas of Stalnaker 1978) and it seems correct.

The disquotation principles introduced by (Kripke 1979) immediately connect subjective information to a theory of *de dicto* belief. According to this theory of belief a sentence describes one of your beliefs if the diagonal determined by the sentence is a superset of your belief alternatives.

This is an alternative approach to the semantics of belief sentences. Certain possible worlds are consistent with all that a person believes. Let us call them belief alternatives. To employ a formulation of Haas-Spohn(1994): if you would put the person into such a possible world she could not discover that it was not the actual world because the world meets everything that she believes. There are no surprises in a belief alternative: they are what the actual world can be according to a belief subject. If we have possible worlds, characters and belief alternatives, we can define a belief operator by stipulating $B_a\varphi$ is true iff for each belief alternative i of a, the proposition expressed by φ at i is true at i[1]. If χ_φ is the character of φ and $bel(a,j)$ are the belief alternatives of a in a world j, we get: $\chi_{B_a\varphi}(i)(j) = 1$ iff $\forall k \in bel(a,j)\chi_\varphi(k)(k) = 1$.

The character based semantics can be transformed in a rather straightforward way into the framework of Hintikka. We define a new set of possible worlds W' from the old worlds W by putting $W' = \{< w,w' >: w,w' \in W\}$. We now define $[\![\alpha]\!]$ as a function from W' to the appropriate α-denotations, by putting $[\![\alpha]\!](< w,w' >) = \chi_\alpha(w)(w')$. The modal accessibility relation (which before was the open relation) becomes $nec(< w,w' >) = \{< w,w'' >: w'' \in W\}$ and the new set of belief alternatives is $bel(a, < w,w' >) = \{< w'',w'' >: w'' \in bel(a,w')\}$. In this way, the belief alternatives are situated on the diagonal of the product and, by the same image, the set of modally accessible worlds from a given world, is the vertical line on which it is situated.

I will further assume that the meaning of non-logical basic expressions is determined in a way that is uniform for their category. That the denotation of John in all possible worlds is the individual John depends on the contingent fact that John bears the name "John" in the actual world. I assume that the principle from which this is an instance is true not just in the actual world, but in all worlds on the diagonal[2]. This principle can mutatis mutandis be extended to (kinds of) verbs, nouns and adjectives. Its full adoption is equivalent to the theory of formal character introduced by Haas-Spohn.

2 Demonstratives and Indexicals Revisited

The theory outlined works fine for *de dicto* belief attributions that are free of variables and indexicals or demonstratives. It is rather easy to show it fails when indexicals occur.

Consider an utterance of (1).

(1) John believes that I am bald.

The diagonal theory forces us into the position that it would be John's opinions as to who the speaker of the sentence is that determines the referent of *I*. Yet John most likely does not have any opinion in that matter as he is not present at the

[1] It may be objected here that we are neglecting an important aspect of character semantics: the fact that contexts give us in addition to possible worlds also speakers, times and places. As this is not an objection to the theory I will arrive at, which does not rely on extra structure in the context fo its treatment of indexicals. Here the reader is invited to add as much structure to the first parameter as she feels is necessary and to live with the consequence of having slightly too much structure in the resulting notion of a possible world.

[2] On a pointed model, we might express this as a schema of MPs like $\forall x(name(\text{“}john''\text{,}x) \leftrightarrow \Box x = john) \wedge \forall y B_y \forall x(name(\text{“}john''\text{,}x) \leftrightarrow \Box x = john)$

occasion of the utterance. We are rather reporting something which John would have rendered (if he in fact believes it) as: He is bald, or That man is bald or even Bill is bald. There is no systematic way of reconstructing what John would have said and thereby it becomes hard to state what *I* contributes to the meaning of the complement.

How to deal with direct reference then? We have found a problem with assuming that direct referentials are "local" elements evaluated by "global" rules. There is a conflict between our analysis of *de dicto* beliefs as holding when they are diagonally true on the belief alternatives of the subject and what happens if I tell you (1) John's absence. John's opinions about who I am are irrelevant. He has a belief about me.

If we assume that belief sentences are diagonally interpreted, the context parameter is the wrong one for indexicals: it equals the context of utterance with one of the belief alternatives of the subject.

We can look for a solution in different directions. For example, we could increase the number of parameters: the fixed context of utterance, the floating context of belief (equal to the first parameter outside beliefs) and the circumstance parameter. We then divide direct referentials into two groups: a group using the first parameter and a group using the second. "Me" would be of the fixed kind, a name would belong to the second.

The problem with this approach is that neither for names nor for indexicals we can uniformly say to what group they belong. Sometimes the belief subject is familiar with the name and its bearer, sometimes he knows the bearer, sometimes he just knows the name. Classically, one would say that in the last case it is a *de dicto* belief, (which means it is non rigid by different values for the second parameter) whereas in the second case the name is interpreted external to the belief. In the first case we should get both: it is interpreted by the first parameter and shifts with the second. So in some cases the name would count as directly referential, in some cases it would be subjective and in some cases finally, it would be both. That the same holds for indexicals is a little bit harder to see. In the example (1) , we obviously have a direct interpretation. But what happens when the belief subject takes part in the current conversation? Can his views about what the indexicals denote be ignored? As it turns out, they cannot but this is better explained after introducing the alternative view.

In the DRT-tradition, a more realistic account is available in the presupposition resolution and accommodation model developed by Van der Sandt. According to this model the treatment of a presupposition trigger (a definite article, a factive verb, a cleft construction etc.) proceeds as follows. We determine at the position of the trigger the content of its presupposition. We then determine if it is possible to resolve the presupposition to material that is accessible from the position of the trigger. Here there is a preference for locally available material. If it is we are nearly done, only the variables in the computed presupposition have to be unified with the corresponding variables in its antecedent. If resolution does not work, we accommodate the material with a preference for the global context, a preference that can be overridden by conflicts with the contents of the global context or with the informativeness of the current utterance.

A DRS can be understood as the structure in which we do the bookkeeping of the emerging common ground between the speaker and the hearer. Normally, this is the bookkeeping of the part that results from linguistic exchanges. Here, we also need to add basic data about the utterance situation: the identity of the speaker, the place

of the utterance, the time and other features. Also, when there is a prior common ground concerning the beliefs of the belief subject or when the belief subject is part of the utterance situation, facts concerning this person should be considered to be part of the DRS in which we are interpreting the sentence.

We also make an extra assumption about presuppositions in belief contexts: if it is not common ground that the subject believes there is nobody who bears the name, we locally accommodate that the name has bearer, also in case the name is externally resolved or accommodated[3].

Let us consider the indexical *I*. In contrast to e.g. the bearer of a proper name, the speaker of the utterance will be part of the common ground between speaker and hearer and it will also be common ground whether or not the belief subject is party to the same common ground. So accommodation instead of resolution does not play any role at all. The identity of the speaker is common ground and it is common ground whether this is also so for the belief subject.

So we get the following predictions concerning an (immediate) occurrence of *I* under a belief operator:

1. If the belief subject is party in the common ground between speaker and hearer (generally this means that she is part of the utterance situation), the discourse referent introduced by *I* will be unified with the discourse referent representing the speaker as the speaker in the beliefs of the belief subject. That discourse referent is identical to the discourse referent representing the speaker in the DRS as a whole.

2. Otherwise, the discourse referent is resolved to the discourse referent that represents the current speaker in the common ground.

This makes a subtle difference in case 1: it rules out that the belief subject will represent me in his belief by a description of which he does not assume that it refers to the same person as the speaker. If I happen to be the last person to leave the building and John does not know this, the fact that John thinks the last person to leave the building is ill, is not sufficient for the truth of the preferred interpretation of (1) .

Why is this theory of indexicals superior? Consider the following situation[4]. Harm

[3]This assumption has been challenged in the otherwise excellent Geurts 1995. Geurts however agrees with the predictions that follow from my assumption. He proposes to derive the predictions by means of a pragmatic mechanism supplementing his own single accommodation or resolution model. This would suit me fine, if he would have such a mechanism on offer. I suspect however that such a mechanism is as hard to find as the mechanism postulated by Karttunen that would derive global accommodations from local ones.

[4]The example

(2) If I would have been the wolf, I would eat you now.

(U. Haas Spohn, p.c.) is another putative counterexample to the theory of direct reference. The example is spoken by the mother goat to her baby goats as she has found the door unlocked against her express instructions. Interpreting the antecedent, we have to resolve *I*. There is in the antecedent no current speaker so the pronoun is resolved to the external speaker, who is in the common ground not just the current speaker and the mother goat but also the person who has just opened the door. The properties of being the current speaker and the mother goat are given up as they are not consistent with the assumption that this is the wolf. So we inherit in the counterfactual situation only: person who just entered the door. To my mind, we do not have genuine counterexample here. "I" still refers to the speaker of the utterance, who in the

has after work made his way to one of the many pubs on the busy street where his office is located. Due to his immodest consumption of beer, he has at one point formed the belief that he is not in the pub in which he happens to be but in the pub on the other side of the street. Looking out of the window he sees in the pub across the street his colleague Wim. I stand next to Harm and seeing him looking as he does and knowing about his state of mind I remark to you:

Harm thinks that Wim is here.

As Harm believes that he is in pub A, which happens to be where he sees Wim, this report is true on the premises of the direct reference theory and also on the widest scope theory about indexicals discussed and rejected by Kaplan[5] But according to my intuitions this interpretation is hardly available.

On our predictions, (2) is false on the preferred interpretation. As Harm shares with us (part of) the speech situation, it is his representation of the place of the utterance (in fact pub B) which prevails and he would not say "Wim is here". After all, the beer has not impaired his vision.

3 What are Individual Concepts?

Let me try to spell out the consequences of these considerations for the meaning of individual terms like demonstratives, first person pronouns and names. Let's assume Harm says *that man* pointing to a balding academic holding a glass of dark beer. What we had to say before is that since Harm is the speaker in the actual world and since he in that world points to a particular person, say the professor of French poetry, it means that person in each of the worlds in the modal hull of the actual world. What we have to say now is more: we have to say what *that man* denotes in the diagonal. The first thing that follows is that it does not always denote something. In some worlds Harm did not say *that man* and in others he omitted the pointing. If we diagonalise over those worlds we fail to get a referent, in the first case because there is nothing that has a relationship to the utterance (it did not take place) in the second case because the rule associated with the pointing does not point: the demonstration is only partially there. The second conclusion is that the meaning is non-rigid. There are certainly worlds where somebody else might have been standing there. Importantly, we may fail to see clearly whom Harm is pointing to and we have thereby belief alternatives in which "that man" has different values.

Much the same holds for the pronoun "I" and for names. The meanings we have to live with for individual terms are non-rigid incomplete individual concepts over our extended set of possible worlds.

Our acceptance of the theory of subjective information has landed us in this predicament and it is an uncomfortable position: we have lost the beautiful and simple theory of quantifying in based on Russell's singular propositions. We are suddenly left —for the attitudes— with partial non-rigid individual concepts and have to make do with them. We have to answer the question which individual concepts there are and how we can deal with them.

counterfactual situation assumed would be the wolf as well (and loses the properties inconsistent with that assumption).

[5]Within our model we would get this theory by ruling out local resolution, i.e. abolishing clause (1).

That not every partial function from (new-style) worlds to individuals is available seems given with our intuitions about belief. We do not want to infer from: *John believes that somebody ate the cake* to *There is somebody such that John believes he ate the cake.* (We combine, for each belief alternative of John, the object that ate the cake there.) We also do not want to give up rigidity, only with extending rigidity to belief alternatives as well. So, we must assume that for each object in the actual world we have a rigid intension over its modal hull. But this assumption offers no help for quantifying into belief contexts, as the intensions so described are insufficient to make any *de re* beliefs true.

My solution is essentially the one that I have advocated in earlier work (Zeevat 1994). That solution attributes to persons objects that are objects of belief. Objects of belief in turn can be counterparts of other objects and can themselves have other counterparts. It is when they are counterparts of real objects, that they satisfy *de re* beliefs. For finding out what objects of belief there are, we have to enquire what experience the subject has had and to what communication she has been exposed.

The difference with my earlier work is that the objects of belief here become partial individual concepts and that we can give an account of their genesis in terms of Hintikka's belief alternatives. I believe the account here improves in some ways on my earlier work. The starting point are cognitive events. These are events in which the subject acquires a new belief. Examples of such events are perceptions, acceptances of an assertion and reasoning. One thing which happens in veridical perceptions and acceptance of true assertions is the genesis of new *de re* beliefs.

Take the example of a perception of an object that the subject has not seen before, let's say Bob's new Ferrari. Before the perception, it will be true that we can describe the Ferrari uniquely by some unique description, say *bob_ferrari.* It then holds before the perception that

$$(3) \qquad \neg\exists x(bob\_ferrari(x) \wedge belief(s, x = x))$$

After the perception this has changed, the subject s has acquired a new object:

$$(4) \qquad \exists x(bob\_ferrari(x) \wedge belief(s, x = x))$$

Of course this is not all that has changed: the belief alternatives of the subject have changed to take in the content of the perception: the beautiful color and the attractive shape will be as the subject perceived them. On the rather unconstructive picture of belief alternatives assumed in this paper the update consists of two steps. First the subject has a unique description of what he sees. The uniqueness follows from the uniqueness of the description *what I am currently seeing* but takes in any other property or relation that the subject attributes to the object in his perception. This first reduces the belief alternatives to those where the description is satisfied (a normal eliminative update.) The second step is the addition to the set of individual concepts of the individual concept formed by the Ferrari (extended to the modal hull of the actual world) and the object that in each of the new belief alternatives is the Ferrari for the subject.

This is the simple case. The object may already be there as the subject may have heard about it before. Two things can happen: the subject recognises the Ferrari from his description so that the updating is limited to a simple update of the alternatives. Alternatively, he does not, which means that the perception will form a second individual concept, which derives from the same Ferrari. The final

possibility is misrecognition: the Ferrari is wrongly recognised as another object the subject is aware of. We get a new concept (the new perception extends the actual Ferrari, coinciding within the belief worlds with the older concept) and an update over the older concept.

Finally, there are hallucinations. Here we only seem to perceive something. It follows that the earlier steps in the updating are as before but the bridge itself cannot be added. We can however add the concept consisting of the referent of the description in each of the belief alternatives.

The general rule which we can abstract for this is that after the perception an individual concept bridges the actual world (if the perception was veridical) and the subject's belief alternatives with the bridge being given by a definite description in the belief alternatives given by the perception itself. Veridical perceptions of an object not seen before can be described as a switch in the truthvalue of a sentence of the form:

$\exists x(x = a \wedge \mathit{believe}(b, x = x))$

Before b's perception the sentence was false, as there was no bridge connecting b's mind to a. After the perception, such a bridge has come into being.

In communication things are not very different, except for the fact that now it is the beliefs of the speaker and the concepts organising his beliefs that take the place of the actual world. I will make the assumption that even in this case it is possible to speak of the object underlying the speaker's utterance: it is the internal object that causes that particular part of the utterance. This assumption can however be eliminated[6]

Bridges will now consist of the speaker's object extended to the belief alternatives of the subject. If the speaker's object was already a bridge it now becomes a longer bridge bringing us to yet other people and possibly to the actual world. The process itself is the same as in the case of perception. There are however two cases where the interpretation process can lead to empty bridges: hallucinated utterances and mishearings are one case, the other is the case of the speaker saying things he does not believe. What we here characterise is the fact that typically communication can lead to a switch in the truth conditions of (5)(e.g. Mary tells Susan about her purported enemy.).

(5) Mary believes that she has an enemy and Susan believes that he is dangerous.

These processes fill the set of individual concepts. The occurence of cognitive events add individual concepts to the given set of individual concepts.

Let us first have a look at the set of individual concepts available in a certain possible world at a given time. This world has a history in which it has acquired new individual concepts. The principle is that either a concept derives from a cognitive event or it was there all the time. We can also say something about what was there all the time. I already mentioned the rigid intensions of the objects of the world. Other concepts that have been always there are necessary objects of belief. I take it that for each subject there is one such object: the self, identified by the subject as the thing that is having these thoughts and experiences. This is a bridge:

[6]The assumption corresponds with the difference between the classical DRT approaches to indefinite reference (no uniqueness) and the approaches based on Evans's work. Non-uniqueness could be accommodated in principle.

it bridges between the subject and what she is in each of her belief alternatives.

Among the individual concepts that exist with respect to a mind there is a distinguished concept that denotes the body of the mind in the actual world and that is defined throughout the belief alternatives of that mind. Originally, the belief alternatives coincide with all that could be known. The self ranges over all these alternatives and realises each possibility as to what it can be. Later as the subject acquires empirical information about itself, certain possibilities with respect to who it can be are eliminated as are the alternatives themselves.

The self is what accounts for *de se* beliefs (Lewis 1979). Adding a property $\mathit{selve}(x)$ to the logical representation language offers solutions to the other problems connected with ego-centric beliefs. Reasons of space prevent any further discussion, here.

Normally, the new individual concepts are (in the set-theoretic sense) extensions of already existing individual concepts, but the idea that all such new concepts form proper extensions is wrong. What happens in a successful communication or in a perception guarantees —if we assume uniqueness— that the originating concept —restricted to the modal hull of the actual world or to the belief alternatives of the sender of the communication — is —when the communication or perception is successful— and the interpreting concept are consistent. If there is overlap between the domains, the uniqueness assumption guarantees that the concept assumes the same values within the overlap. This is however not enough for unique extendibility. A simple example may suffice. Assume that on my way home from work, I have witnessed a traffic accident. Coming home, I tell my wife about the accident in the certain opinion that she will not tell our daughter about this, as she prefers to keep our daughter from such harsh realities. Yet, the next day my daughter tells me about an accident that happened. I assume that she is speaking of another accident, and not that my wife has talked. But my wife has and she is talking about the same accident. My concept of the accident is inconsistent with the concept that causes it, which contains my concept of the original accident.

To solve this problem we could make a formal distinction between the actual world and between the belief alternatives of each subject. This can be done using the selves. Another solution is to switch to a notion of semi-extension with respect to a belief subject. A semi-extension with respect to me, can overwrite values for the concept in my belief alternatives. As the causing concept will continue to exist, they can be recovered. In the example, it holds that there is an accident of which I believe conflicting things (e.g. that my daughter knows about it and that she does not know about it) but it is surely wrong to attribute to me an internal inconsistency). I will in the sequel, ignore this problem.

4 Some Refinements

Let us consider the Hob-Nob-sentence (6).

(6) Hob believes that a witch poisoned his cow and Nob believes that she killed his pig.

On the hardest interpretation of the sentence there is a rumour in the village about a witch and Hob and Nob are parties to that same rumour (due to a conflict they never talk with each other). In terms of our model, this means that there are two

chains to the originator of the rumour, but not chains such that one extends the other. The sentence is true therefore if there is a chain (of which neither Hob or Nob is the endpoint) which can be extended to both the Hob chain and the Nob chain.

We can write this in the following way:

(7) $\exists xyz(x \subseteq y \wedge x \subseteq z \wedge$
$belief(hob, witch(y) \wedge Py) \wedge belief(nob, Qy))$

There are two problems with this representation. First, it seems we have a *de re*-version of an obviously *de dicto*-sentence. and, second, the sentence has one quantification and not three.

Let us start with the second problem, which is less serious anyway. We can build into the semantics of the belief operator the possibility to extend individual concepts. This is easily implemented.

(8) $M, i, g \models belief(x, \varphi) \Leftrightarrow \exists h(\forall v \in VAR\ (g(v) \subseteq h(v)) \wedge \forall j \in bel_i(g(x))M, j, h \models \varphi)$
$M, i, g \models \neg belief(x, \varphi) \Leftrightarrow \forall h(\forall v \in VAR\ g(v) \subseteq h(v)) \rightarrow \exists j \in bel_i(g(x))M, j, h \models \neg\varphi$

Using this semantics we can reduce our representation to (9).

(9) $\exists x(belief(hob, witch(x) \wedge Px) \wedge belief(nob, Qx))$

The first problem can be solved by undoing the existential import of the existential quantification. In order to be able to express existential import, we have an existence predicate with the semantics as in (10).

(10) $M, i, g \models Ex \Leftrightarrow g(x)(i)$ defined
$M, i, g \models \neg Ex \Leftrightarrow g(x)(i)$ is undefined

The normal existential quantifier representing the "there is" of natural language will be now rendered by $\exists x(Ex \wedge \ldots)$. Existential quantification is defined by (11).

(11) $M, i, g \models \exists x\varphi \Leftrightarrow$ there is a $d \in IC_i$ such that $M, i, g[d/x] \models \varphi$
$M, i, g \models \neg\exists x\varphi \Leftrightarrow \forall d \in IC_i M, i, g[d/x] \models \neg\varphi$

With our E-predicate we can represent the *de dicto*/*de re*-distinction as in (12).

(12) $\exists x belief(y, \varphi)$
$\exists x(Ex \wedge belief(y, \varphi))$

This leads to a simple treatment of the Hob-Nob sentence.

The final problem we have to face is how to link this representation to syntax. The simplest solution is to call on DRT once more. We need to add the proviso that a belief box does not have place for its own discourse referents: they drift away to the embedding box. We have to distinguish them however from other discourse

referents. This can be achieved by making sure that all other discourse referents are entered with a condition ensuring their existence. This is superfluous in case the expression responsible for their introduction also places a condition in which they appear as an argument in the box: the standard semantics for simplex conditions lets them entail the existence of all their arguments, e.g. from $give(x, y, z)$ we can infer Ex, Ey and Ez. But foreign discourse referents (the ones not also occurring in an argument position in a simplex condition in their box) do not need to exist. This can sometimes make it necessary to bring in a condition Ex in the other rules of the development algorithm.

Is there something wrong with representing *de dicto* belief reports as we have? I think not, in fact it is completely the reverse: there are serious problems with the traditional representation. On the picture I have been developing individual concepts also arise when it only appears to the subject that an object causes a perception or assertion. The concept will then be restricted to the belief alternatives of the subject. In traditional terminology, an internal object is created for which there is no external object. The traditional representation of *de dicto* existential sentences completely misrepresents this fact. For a *de dicto* existential to be true on the standard representation, it is sufficient that the non-satisfaction of a predicate is not consistent with whatever the subject believes.

But intuitively this not the same as the subject believing the existential sentence. Belief requires that she has noticed the incompatibility of the non-satisfaction of the predicate with her other beliefs which on the classical view would be equivalent to postulating an object satisfying the predicate. This postulation is a cognitive act, comparable to perception and believing an assertion, and it would follow that the object must be an individual concept. (On the intuitionist view, it would not even be possible to attribute an existential *de dicto* belief on the basis of the inconsistency of the non-satisfiability of the predicate with the subject's other belief.)

Our account of the emergence of individual concepts faithfully captures these reasoning processes: coming to believe in an existential sentence is the creation of an individual concept that is minimally defined over the belief alternatives of the subject.

5 Proper Names

The treatment of proper names is an important test case in semantical theories of belief. Rather than starting from the premiss that a name is an individual term with an autonomous meaning —the assumption in most of the philosophical literature— I will start from the opposite point of view, namely that names are a special kind of anaphoric device, much like anaphoric pronouns. This starting point makes it possible to explain some of the properties of names that do not come into the picture if we follow the philosophical literature, such as their presuppositional properties and the possibility of using the same name for different people in the same context. Haas-Spohn's 1994 shows some of the difficulties that are encountered when we start from the other premiss. Let me point out in addition that the anaphoric theory is the only practical option in natural language interfaces for the computer.

Names and anaphoric devices such as personal pronouns and definite descriptions have in common that they are all devices that find a suitable antecedent in the common ground and pass the value that they found to the construction at hand. Names and pronouns of course differ in the way in which they do this. Personal

Jaap van der Does jvddoes@fwi.uva.nl
Department of Mathematics and Computer Science, University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
The Netherlands

Markus Egg egg@heidelbg.ibm.com
Institut für Logik und Linguistik, IBM Deutschland GmbH
Postfach 10 30 68
D-69020 Heidelberg
Germany

Martin Emms emms@cis.uni-muenchen.de
Centrum für Informations- und Sprachverarbeitung, Ludwig Maximilians Universität
Wagmüllerstraße 23
München 80538
Germany

Tim Fernando fernando@ims.uni-stuttgart.de
IMS, Universität Stuttgart
Azenbergstraße 12
D-70174 Stuttgart
Germany

Tsutomu Fujinami tsutomu@ims.uni-stuttgart.de
Institut für Maschinelle Sprachverarbeitung, Stuttgart
Azenbergstraße 12
D-70174, Stuttgart
Germany

Brendan S. Gillon gillon@langs.lan.mcgill.ca
Department of Linguistics, McGill University
1001 Sherbrooke Street West
Montreal Quebec
H3A 1G5 Canada

Sheila Glasbey cavedon@cs.rmit.edu.au
Centre for Cognitive Science, University of Edinburgh
2 Buccleuch Place
Edinburgh EH8 9LW
United Kingdom

Willem Groeneveld groenev@illc.uva.nl
ILLC/Department of Philosophy, University of Amsterdam
Nieuwe Doelenstraat 15
1012 CP Amsterdam
The Netherlands

Javier Gutierrez Rexach izzy362@mvs.oac.ucla.edu
Department of Linguistics, UCLA
405 Hilgard Avenue
Los Angeles, CA 90024–1543
USA

Herman Hendriks herman.hendriks@let.ruu.nl
Research Institute for Language and Speech, Utrecht University
Trans 10
3512 JK
Utrecht
The Netherlands

Marco Hollenberg hollenb@phil.ruu.nl
Department of Philosophy, University of Utrecht
Heidelberglaan 8
3584 CS Utrecht
The Netherlands

Antoon Hurkmans ---
Vrije Universiteit Amsterdam
Koningsstraat 46-I
1011 EW AMsterdam

Gerhard Jäger jaeger@asg.ag-berlin.mpg.de
Arbeitsgruppe SG der MPG, CIS München
Jägerstraße

Theo M.V. Janssen theo@fwi.uva.nl
ILLC / Fac. WINS, University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam

Jan Jaspars jaspars@cwi.nl
CWI, Amsterdam
P.O.Box 94079
1090 GB Amsterdam

László Kálmán kalman@nytud.hu
Theoretical Linguistics Programme, Budapest University
Research Institute for Linguistics, Room 119
P.O.Box 19
H-1250 Budapest
Hungary

Megumi Kameyama megumi@ai.sri.com
AI Centre, SRI International
333 Ravenswood Avenue
Menlo Park
CA 94025
U.S.A.

Michael Kohlhase kohlhase@cs.uni-sb.de
Computer Science, Universität des Saarlandes
Postfach 1150
D-66041 Saarbrücken
Germany

Emiel Krahmer krahmer@natlab.research.philips.com
Institute for Perception Research, Eindhoven
P.O. Box 513
5600 MB Eindhoven
The Netherlands

Susanna Kuschert kuschert@coli.uni-sb.de
Computational Linguistics, Universität des Saarlandes
Postfach 1150
D-66041 Saarbrücken
Germany

David Milward milward@coli.uni-sb.de
Computerlinguistik, Universität des Saarlandes
Postfach 1150
D-66041 Saarbrücken
Germany

Michael Moortgat moortgat@let.ruu.nl
Research Institute for Language and Speech, Utrecht University
Trans 10
3512 JK Utrecht

Larry Moss lsm@cs.indiana.edu
Departments of Mathematics and Computer Science, Indiana University
Bloomington IN 47405 USA

Seungho Nam nam@power2.snu.ac.kr
Department of Linguistics, Seoul National University
Sillim-Dong, Kwanak-ku
Seoul [151-742], Korea

Ralf Naumann naumann@ling.uni-duesseldorf.de
Seminar für Allgemeine Sprachwissenschaft, Universität Düsseldorf
Universitätsstraße 1
40225 Düsseldorf
Germany

Manfred Pinkal pinkal@coli.uni-sb.de
Computational Linguistics, Universität des Saarlandes
Postfach 1150
D-66041 Saarbrücken
Germany

Massimo Poesio poesio@cogsci.ed.ac.uk
Centre for Cognitive Science, University of Edinburgh
2 Buccleuch Place
Edinburgh EH8 9LW
Scotland UK

Livia Polanyi livia@csli.stanford.edu
CSLI Stanford, and Rice University
Ventura Hall
Stanford CA 94305-4115
U.S.A.

Ekaterina V. Rakhilina plungian@iling.msk.su
VINITI, Russian Academy of Sciences
Medvedeva 8–83
103050 Moscow
Russia

Gene Rohrbaugh rohrbaugh@mail.utexas.edu
Department of Linguistics, University of Texas at Austin
Austin TX 78712
USA

Hotze Rullmann rullmann@let.rug.nl
Vakgroep Nederlands, Rijksuniversiteit Groningen
Postbus 716
9700 AS Groningen
The Netherlands

Anna Szabolcsi ibaesza@mvs.oac.ucla.edu
Linguistics, UCLA
405 Hilgard Avenue
Los Angeles CA 90024-1543
U.S.A.

Matt Watson mwatson@la.utexas.edu
Philosophy Department, University of Texas at Austin
Waggener Hall
Austin TX 78722
USA

Robert R. Westmoreland rwestmor@indiana.edu
Department of Linguistics, Indiana University
322 Memorial Hall, Indiana University
Bloomington
Indiana 47404
USA

Yoad Winter yoad.winter@let.ruu.nl
OTS, Utrecht University
Trans 10
3512 JK Utrecht
The Netherlands

Henk Zeevat henk@mars.let.uva.nl
Department of Computational Linguistics, University of Amsterdam
Spuistraat 134
1012 VB Amsterdam