Relational Semantics and Scope Disambiguation

Massimo Poesio
Department of Computer Science, University of Rochester
Rochester, NY 14627, USA[*]

Abstract

Relational Semantics can be used to give a denotation to the unscoped logical forms used by Natural Language Processing systems - representations in which the quantifiers are 'left in place' rather than extracted. In this way, the system is no longer required to compute all the disambiguated interpretations of a sentence before storing its representation in the knowledge base. Rules of inference can also be defined so that the disambiguation process can be modeled by formal derivations, and 'weaker' versions of the standard rules of inference can be specified so that conclusions can be drawn without requiring a complete disambiguation.

1 Introduction

In using (1), a speaker could mean that there is some one undergraduate that is dating all male students, or merely that all male students date some undergraduate or other.

(1) Every male student dates an undergrad.

The conventional view is that each sentence with multiple interpretations is to be seen as ambiguous, that is, each interpretation has to be represented by a distinct formula. The two interpretations of (1) are represented by (2a) and (2b).

(2) a. $((\forall x)\ (\mathrm{MS}(x) \supset ((\exists y)\ \mathrm{U}(y) \wedge \mathrm{D}(x,y))))$
b. $((\exists y)\ (\mathrm{U}(y) \wedge ((\forall x)\ (\mathrm{MS}(x) \supset \mathrm{D}(x,y)))))$

'Traditional'[1] Natural Language Processing (NLP) systems, such as TEAM [Grosz *et al.*, 1987] or the Core Language Engine [Alshawi *et al.*, 1988], are built according to this view, and analyze (1) more or less as follows. First, the parser computes a *logical form* [Webber, 1978; Schubert and Pelletier, 1982; Allen, 1987; Alshawi and van Eijck, 1989] which is similar to the S-structure representation of (1) before Quantifier Raising:

(3) [<every x male-student> dates <a y undergrad>]

All the unambiguous interpretations of (1) are then extracted from (3) by algorithms like that proposed by Hobbs and Shieber [1987]. Finally, the system has to choose an interpretation, which is normally done using preference heuristics [Hurum, 1988].

[*]This paper appears in Barwise, J. and Gawron, J. M. and Plotkin, G. and Tutiya, S. (editors), Situation Theory and its Applications, vol. 2, 1991, CSLI, p.469–498. The work was in part supported by the ONR/DARPA research contract number N00014-82-K-0193 and the NSF research grant number CDA-8822724.

[1]The use of 'traditional' here should not be thought of as derogatory. Actually, it is more synonymous with 'working'.

The disadvantages of this way of doing things have not gone unnoticed [Kempson and Cormack, 1981; Hobbs, 1982; Allen, 1991]. In this paper I will especially focus on two problems. The first problem is that a system built in this way cannot use the information which comes later in the discourse. Yet, this information could save the system a lot of work. Suppose, for instance, that sentence (1) is followed by sentence (4); one could immediately conclude that *an undergrad* in (1) has to take wide scope.

(4) I met her yesterday.

The second problem is that because of the great number of possible interpretations, computing all of them can be very expensive. Two factors combine to multiply the number of interpretations. First, the number of (scopally) distinct interpretations grows as the factorial of the number of NPs. Sentence (5) has $5! = 120$ interpretations, all distinct (in the sense that no two interpretations are such that one entails the other).

(5) In most democratic countries most politicians can fool most of the people on almost every issue most of the time. [Hobbs, 1983]

Yet, it seems obvious that people do not entertain 120 possibilities when hearing (5b) In addition, scope is just one of the reasons why a sentence may have more than one interpretation. By just considering the fact that in a sentence with numerical quantifiers like (6a) the examiners may be involved to a different degree in the grading, Kempson and Cormack [1981] are able to find at least four interpretations for it:

(6) a. Two examiners marked six scripts.
b. Three Frenchmen visited five Russians.

(6a) can be used to mean (a) that the same six scripts were each marked by two examiners, (b) that two examiners marked six (not necessarily the same) scripts each, (c) that two examiners marked a group of six scripts between them, and (d) that two examiners each marked the same set of six scripts. As for (6b), Partee ([1975], quoted by Bunt [1985]) argues that it has eight readings; Bunt, counting also collective and distributive interpretations, is able to find 30 different readings for it! Even 30 interpretations seems to be too large a number to be actually considered. And yet, people do understand (5b) and (6b).

The aim of this article is to show that a fairly simple solution to the two problems just presented is possible within the 'relational' semantic framework developed in [Barwise and Perry, 1983; Barwise, 1987; Rooth, 1987]. The basic idea of relational semantics is that certain expressions 'change' the conditions of interpretation, represented by the variable assignments. For example, after the sentence *A man came in* has been added to a discourse, every variable assignment f which satisfies the new discourse has to assign a value to the variable used to represent the NP *a man*. This process can be modeled most effectively by using partial variable assignments and by requiring that each variable assignment which satisfies the whole discourse be an extension of a variable assignment which satisfies the portion of discourse prior to the last sentence. The value of a sentence will not be a truth value, but a *relation*, that is, a set of pairs of assignments, where the first element of each pair will be a variable assignment which satisfies the discourse prior to the addition of the expression, and the second element will be a variable assignment which satisfies the discourse after the sentence has been added [Barwise, 1987; Rooth, 1987].

My solution will be to give to the 'unscoped' expressions proposed in the NLP literature a denotation which is the union of the denotations of the disambiguated interpretations, and to use semantically justified inference rules to derive from the unscoped representation a new representation which is consistent with a smaller number of interpretations, possibly just one.[2] A system which uses a representation like the one I will present does not have to compute all the interpretations of (1) right away. Once such a system has translated (1) into an unscoped logical form, this logical form can be immediately stored in the discourse representation. As a matter of fact, it is not necessary for the system to disambiguate at all, unless required to do so: I will also show rules to derive certain types of conclusions from an unscoped representation. If a sentence like (4) is also asserted, however, and if the system is able to conclude that *her* in the second sentence is anaphoric to *an undergrad* in the first one, it will also be able to conclude that *an undergrad* takes scope over *every male student*. It should be clear that in this way both problems with the previous kind of architecture are solved.

I will use a DRT style of representation of the kind proposed in [Kamp, 1981]. I needed a discourse representation which would make it easier to talk about anaphoric relations, and I chose DRT with the hope that this kind of representation would be better known than, say, Dynamic Predicate Logic [Groenendijk and Stokhof, 1990] or Episodic Logic [Schubert and Pelletier, 1988]. However, I will redefine Kamp's semantics for DRT, and in such a way that the difference between the different kinds of representations will reduce considerably.

The organization of the paper is as follows. I will first discuss three solutions for the two problems of using later information and having to compute all the interpretations: the idea of using disjunction, the 'radical vagueness proposal' of Kempson and Cormack, and Hobbs' solution based on 'dependency functions'. I will explain why these solutions are all incomplete in one way or the other. After having introduced my DRT notation, and having shown how one can write a relational semantics for it preserving the properties of the representation proposed in [Kamp, 1981], I will show in section 4 that one can use this semantics to add to the logic a construct to represent scopally ambiguous sentences, the *scope forest*, as well as to give a semantic justification to the disambiguating inference rules. Before introducing the logic more formally, I will give examples of 'disambiguation by deduction' in section 5. I will close with a discussion of some issues raised by the solution I propose.

2 Previous Work

I am aware of three kinds of solutions to the two problems with the conventional view. In this section I will review them in turn.

2.1 Disjunction

The easiest way for representing a sentence with multiple interpretations without losing any information is to represent that sentence with a disjunction of all the interpretations. Sentence (1), for example, would be represented by the disjunction (7).

(7) $((\forall x)\ (\textsc{ms}(x) \supset ((\exists y)\ \textsc{u}(y) \wedge \textsc{d}(x,y)))) \ \vee\ ((\exists y)\ (\textsc{u}(y) \wedge ((\forall x)\ (\textsc{ms}(x) \supset \textsc{d}(x,y)))))$

In this way, it is possible to take advantage of disambiguating information that may come later. This solution has several problems, however. First of all, as Kempson and Cormack point out, is

[2] To dissipate possible misunderstandings, it may be useful to add that even though the logic I will present is strictly monotonic, I'm not claiming that discourse disambiguation is a monotonic process. Some of the aspects of discourse inference can be explained as the result of the interaction of a number of different constraints, and at the moment I am mostly interested in exploring these interactions. Nothing however prevents extending the treatment I'll propose with nonmonotonic or probabilistic rules. I will return on this topic at the end of the paper.

the *Mapping Problem* - there are a number of reasons for preferring as the semantic representation of a sentence a logical structure as close as possible to its syntactic structure, which isn't the case for (7). A second, and I think decisive, argument against this method is that it requires all the interpretations to be computed, and therefore is not a solution to the combinatorial explosion problem[3]

2.2 Vagueness

Kempson and Cormack [1981] contend that the conventional view is misled. Even if (1) or (6a) have different interpretations (as K&C put it, they are *logically* ambiguous), they claim that those sentences are not *linguistically* ambiguous, that is, they have a single semantic representation. In their view, the representation of a sentence with multiple interpretations is the weakest representation entailed by all interpretations. This proposal works fairly well for sentences like (1), because in fact the two interpretations of that sentence are not really distinct: the reading in which a single undergraduate is dating all male students entails the other. The representation initially proposed by Kempson and Cormack for (1) is (8).

(8) $\exists M \; \forall m_{m \in M} \; \exists U_1 \; \exists u_{u \in U} \; \mathrm{D}(m, u)$

In order to extend this method to represent sentences like (6a), however, something more drastic is called for, since none of the interpretations of (6a) is entailed by all other three; the 120 interpretations of (5b) are also all distinct. In order to give a unique semantic representation to sentences like these, Kempson and Cormack have to present a second version of the theory in which a much weaker representation is used. The representation of (6a) is (9a), and the new representation of (1) is shown in (9b).

(9) a. $\exists X_2 \; \exists S_6 \; \exists x_{x \in X_2} \; \exists s_{s \in S_6} \; \mathrm{M}(x, s)$
b. $\exists M \; \exists U_1 \; \exists m_{m \in M} \; \exists u_{u \in U} \; \mathrm{D}(m, u)$

All that (9b) says is that there are a set of male students and a set of undergrads, and that one male student dated one undergrad. These truth conditions are of course much too weak a representation of (1): a NLP system using (9b) as the representation for (1) would have to pay a high price to eliminate the problems discussed above.

This is not, however, what Kempson and Cormack have in mind. Their idea is that (9b) is not the final representation of (1), but only the 'basis' from which the real interpretations can be generated by means of two operations: *uniformising* - when an existential quantifier follows a universal, reverse their order - and *generalising* - turn an existential quantifier into a universal. But if this is what they have in mind, we are left with something not much different from a 'traditional' system - the extraction operations do the job that in a traditional system would be done by an algorithm like Hobbs and Shieber's (assuming of course than one can justify these operations semantically, which Kempson and Cormack don't) and the 'filters' that Kempson and Cormack use to choose one interpretation are not much different from the preference heuristics mentioned above.

2.3 Dependency between sets

The proposal advanced by Hobbs in [Hobbs, 1983] falls in a third class of solutions, based on finding dependencies between sets. Hobbs' solution is based on a set of assumptions. First, Hobbs wants to use a first-order representation, with variables ranging over sets. Second, he represents determiners

[3] Finding the logical representation for (5) is left as an exercise to the reader.

as relations between sets. The sets he has in mind are not, however, the set of sets denoted by the NP and the set denoted by the VP. For example, he paraphrases a sentence like *Most men work* as 'there exists a set s which represents a majority of the set of all men, and for each individual y in s, y works'. This paraphrase becomes in his representation the formula (10).

(10) $(\exists s)\ (\text{MOST}(s, \{x|\text{MAN}(x)\}) \wedge (\forall y)(y \in s \supset \text{WORK}(y)))$

Hobbs' third assumption is that sets have *typical elements*. The typical element of a set s is an individual $\tau(s)$ defined by the following axiom:

(11) $(\forall s)P_s(\tau(s)) \equiv (\forall y)(y \in s \supset P(y))$

Where P_s is a predicate which is like P except that it is also true of $\tau(s)$ iff P is true of all the elements of s. Hobbs' representation for (1) is (something like) (12), which can be read as follows: there is a set m which includes all the male students, a set u which contains one undergrad, and the typical element of m dates the typical element of u.

(12) $(\exists m, m_1, u, u_1)\ (\text{EVERY}(m, m_1) \wedge\ \text{A}(u, u_1) \wedge\ \text{MALE-STUDENT}_{m_1}(\tau(m_1)) \wedge$ $\text{UNDERGRAD}_{u_1}(\tau(u_1)) \wedge \text{DATES}(\tau(m), \tau(u)))$

Finally, scope relations are represented using *dependency functions*. A dependency function f returns, for each male student x, the set of undergrads that x dates:

$$f(x) = \{y|\ \text{UNDERGRAD}(y) \wedge \text{DATES}(x, y)\}.$$

If the inferencing component discovers there is a different set u for each element of the set m, u can be viewed as referring to the typical element of this set of sets, and the fact $u = \tau(\{f(x)|x \in m\})$ can be added to the knowledge base. There are two problems with this solution. First of all, as Hobbs points out, the representation in (10) can only be used with monotone increasing determiners, like *most* and *every*. For, if we were to represent *No man works hard* in Hobbs's representation, we would be able to conclude that no man works, which doesn't follow - *no* is not monotone increasing. The second problem, common to other dependency function-based solutions, is that only sentences with two quantifiers can be given a scope-neutral information, and not, say, sentences with a quantifier and negation, as *John doesn't have a car.*

3 A Relational Semantics for DRT

This section has two purposes: introducing the DRT-based representation I will use in the rest of the paper, together with some terminology[4], and show how one can specify a 'relational' semantics for such a representation. I will call the resulting logic DRT_0. The reader to whom both DRT and relational semantics are familiar may want to skip this section.

3.1 Syntax of DRT_0

The set of symbols of DRT_0 includes a set of **property symbols** (unary predicates), a set of **relational symbols**, and a set of **markers**: $x_0, \ldots, x_n, \ldots$. (I will sometimes use for simplicity letters without indices like x, y, etc. for the markers.) The set of expressions of DRT_0 consists of:

1. **marker introducers** like α^{x_i}, where x_i is a *new* marker, that is, a marker not used for any other marker introducer. (That is, i must be strictly greater than any previously used marker index.)

[4]The amount of space at my disposal prevents a complete introduction to DRT.

2. **conditions**:

   (a) **unary conditions** like $P(x_i)$, where x_i is a marker and P is a property symbol.

   (b) **binary conditions** like $R(x_i, x_j)$, where x_i and x_j are markers and R is a relation symbol.

   (c) **coindexing conditions** like $x_i = x_j$, where x_i and x_j are markers.

   (d) **negated DRSs** of the form $\neg K$, where K is a DRS.

   (e) **conditional DRSs** of the form $K_1 \rightarrow K_2$, where K_1 and K_2 are DRSs.

3. **Discourse Representation Structures**: a DRS is an expression containing one or more conditions and zero or more marker introducers, usually written as in (13), where α^{x_0} and α^{x_1} are marker introducers, FARMER(x_0) and DONKEY(x_1) are unary conditions, and OWNS(x_0,x_1) is a binary condition.

$$(13)\quad \boxed{\begin{array}{l} \alpha^{x_0}\ \ \alpha^{x_1} \\ \text{FARMER}(x_0) \\ \text{DONKEY}(x_1) \\ \text{OWNS}\ (x_0,x_1) \end{array}}$$

In keeping with the standard conventions, I will reserve the symbol K, possibly with subscripts, to indicate DRS's. Subscripted x's like x_0 will always indicate markers. Let a marker x be **free in** K if no marker introducer α^x is in K. A DRT$_0$ **formula** is a DRS with no free markers. The donkey sentence *Every farmer who owns a donkey beats it* is represented in DRT$_0$ by (14).

$$(14)\quad \boxed{\boxed{\begin{array}{l} \alpha^{x_0}\ \ \alpha^{x_1} \\ \text{FARMER}(x_0) \\ \text{DONKEY}(x_1) \\ \text{OWNS}\ (x_0,x_1) \end{array}} \rightarrow \boxed{\begin{array}{l} \alpha^{x_2} \\ x_2 = x_1 \\ \text{BEATS}(x_0,x_2) \end{array}}}$$

The only significant difference between DRT$_0$ and 'standard' DRT is the distinction between 'use' and 'introduction' of markers. This distinction makes it easier to enforce the constraint that each marker has to be new, as well as simplifying the definition of the semantics of a DRS, but otherwise has no semantic consequences.

3.2 The Semantics

A model M for DRT$_0$ is a pair $\langle U, F\rangle$: U is a nonempty set, and F an interpretation function. Assignments are called **embedding functions** in DRT; embedding functions are partial functions from markers to objects of the domain. An embedding function over M is a function which associates to the markers values from U. The denotation with respect to M of an expression of DRT$_0$ is a set of pairs of embeddings over M defined as follows:

1. $\|\alpha^{x_i}\|^M = \{\langle f, g\rangle \mid f \subseteq g,\ x_i \notin \text{DOM}(f) \text{ and } g = f\ \cup \langle x_i, a\rangle, \text{ for some } a \in U\}$

2. $\|\text{FARMER}(x_i)\|^M = \{\langle f, f\rangle \mid f(x_i) \in F(\text{FARMER})\}$

3. $\|\text{OWNS}(x_i, x_j)\|^M = \{\langle f, f\rangle \mid \langle f(x_i), f(x_j)\rangle \in F(\text{OWNS})\}$

4. $\|x_i = x_j\|^M = \{\langle f, f\rangle \mid f(x_i) = f(x_j)\}$

5. $\left\| \begin{array}{|l|}\hline \alpha^{x_1}, \ldots, \alpha^{x_n} \\ C_1 \\ \vdots \\ C_m \\ \hline \end{array} \right\|^M = \{\langle f, f\rangle |$ there exist $f_1 \ldots f_n$ $\langle f, f_1\rangle \in \|\alpha^{x_1}\|^M, \ldots \langle f_{n-1}, f_n\rangle \in \|\alpha^{x_n}\|^M$ and $\langle f_n, f_n\rangle \in \|C_1\|^M, \ldots \|C_m\|^M\}$

6. $\|\neg K\|^M = \{\langle f, f\rangle | \langle f, f\rangle \notin \|K\|^M\}$

7. $\|K_1 \rightarrow K_2\|^M = \{\langle f, f\rangle |$ for all f s.t. for all extensions g s.t. $\langle f, g\rangle \in \|K_1\|^M$, there exists h s.t. $\langle g, h\rangle \in \|K_2\|^M\}$

It is easy to check that the verification conditions in DRT_0 are analogous to those of standard DRT, and to verify that by requiring that a marker x can only be coindexed with a marker y if the assignment x is defined on both, one also obtains the same accessibility conditions of DRT. Truth can be defined as follows. A formula K is **true** in a model M iff $\|K\|^M \neq \emptyset$. A simple notion of entailment for DRT_0 can be defined as follows: if K_1 and K_2 are formulas, $K_1 \models K_2$ iff for all models M in which K_1 is true, $\|K_1\|^M \subset \|K_2\|^M$.

3.3 Inference Rules

I am not aware of any definition of inference rules for DRT in the literature, so I will introduce one that will do for the purposes of this article. A *rule of inference* in DRT_0 is a way of deriving a *conclusion* from a set of *premises*, precisely as in first order logic. That it, the rules of inference of DRT are of the form

$$\frac{K_1, \ldots, K_n}{K}$$

where both the premises $K_1, \ldots, K_n$ and the conclusion K are conditions. The only difference is that these inference rules will be DRS-specific, in the sense that the argument is applicable only when the premises $K_1, \ldots, K_n$ are all conditions of a single *formula* K'; the conclusion will also be added to the same DRS, obtaining a new DRS K''. An inference rule is **acceptable** iff K'' is still a formula. A rule of inference will be **sound** iff it is acceptable, and $K' \models K''$. An example of sound rule of inference for DRT_0 is the following version of *Modus Ponens*:

$$\frac{\boxed{\mathrm{P}} \longrightarrow \boxed{\mathrm{Q}},\ \mathrm{P}}{\mathrm{Q}}$$

4 The Proposal

DRT_0 is not a solution to our problems: the two interpretations of sentence (1), in fact, still have to be represented by distinct DRT_0 formulas:

(15)

$$a.\ \begin{array}{|c|}\hline \begin{array}{|l|}\hline \alpha^x \\ \textsc{male-student}(x) \\ \hline \end{array} \longrightarrow \begin{array}{|l|}\hline \alpha^y \\ \textsc{undergrad}(\mathrm{y}) \\ \textsc{dates}(x, y) \\ \hline \end{array} \\ \hline \end{array}$$

$$b.\ \begin{array}{|l|}\hline \alpha^y \\ \textsc{undergrad}(\mathrm{y}) \\ \begin{array}{|l|}\hline \alpha^x \\ \textsc{male-student}(x) \\ \hline \end{array} \longrightarrow \begin{array}{|l|}\hline \textsc{dates}(x, y) \\ \hline \end{array} \\ \hline \end{array}$$

Because of the way the semantics has been defined in the previous section, however, it will be relatively easy to extend DRT_0 with a new construct which will be used to give a unique representation to (1). In this section I will introduce a model of disambiguation in which sentences are represented as **scope forests** whose denotation is the union of the denotations of the scopally disambiguated interpretations, and the number of possible interpretations can be restricted using inference rules which reflect either logical or referential facts.

4.1 Scope Forests

Consider first a slightly modified version of (1).

(16) Every male student dates most undergrads.

The 'unscoped logical form' of (16) proposed in the literature [Schubert and Pelletier, 1982; Allen, 1987; Alshawi and van Eijck, 1989] has a form that in DRT terms could be something like (17).

(17) $$\boxed{< \mathbf{every}_1 \boxed{\begin{array}{l}\alpha^x \\ \text{MALE-STUDENT}(x)\end{array}} > \text{DATES} < \mathbf{most}_2 \boxed{\begin{array}{l}\alpha^y \\ \text{UNDERGRAD}(y)\end{array}} >}$$

My goal is to give a denotation to a DRS of this type, and the relational semantics sketched in section 3 can be used to this end, as follows. There are two quantified NPs in the sentence, *every male student* and *most undergrads*, and two ways of 'ordering' them to get an interpretation. The interpretation of (16) in which the universal takes wide scope (assuming a representation for generalized quantifiers in DRT roughly analogous to the one Kamp suggests in [1988]) is shown in (18).

(18) $$\boxed{\boxed{\begin{array}{l}\alpha^x \\ \text{MALE-STUDENT}(x)\end{array}} \xrightarrow{\mathbf{every}_1} \boxed{\boxed{\begin{array}{l}\alpha^y \\ \text{UNDERGRAD}(y)\end{array}} \xrightarrow{\mathbf{most}_2} \boxed{\text{DATES}(x,y)}}}$$

If we think of the restrictions of the quantifiers and of the DRS which represents the scope of *most undergrads* in (18) as nodes of a graph, we can see each way of ordering the quantifiers as a path which starts from the antecedent of the quantifier with wider scope and ends with the consequent of the quantifier with narrower scope. I will therefore call each such way of ordering the NPs a **path**. In the relational semantics of section 3, each path denotes a set of pairs. The denotation of the logical form in (17) can therefore be defined as the union of these sets. I will call DRS's like (17) **scope forests**. If the denotation of a scope forest *SF* is the union of the denotation of the paths in a set *PS*, I will say that the paths in *PS* are **associated** with *SF*. The translation rules in the grammar will be such that the interpretation of a sentence like (16) is a scope forest[5].

4.2 Ordering Constraints and Scope Forest Disambiguation

As new facts about the relative scope of the NPs in a sentence S are discovered, the number of paths (that is, interpretations) associated with the scope forest *SF* representing S is reduced. The inference rules for scope disambiguation (below) model this process. These rules derive from a scope forest *SF* a new scope forest *SF'* which has more **ordering constraints** (using logical truths, facts about reference, etc.) . Each ordering constraint of a scope forest *SF* is a label of the form $i < j$,

[5]The logical form is generated as proposed by Schubert and Pelletier [1982]. The fragment cannot be included in the paper, because of space constraints.

where i and j are indices of operators in SF, and indicates that only paths in which the operator with index i precedes the operator with index j are associated to SF. For example, the scope forest equivalent to the subset of interpretations of (16) in which *most undergrads* takes scope over *every male student* is represented by the DRS in (19).

(19)
$$\varphi \boxed{< \mathbf{every}_1 \boxed{\begin{array}{l}\alpha^x \\ \text{MALE-STUDENT}(x)\end{array}} > \ \text{DATES} \ < \mathbf{most}_2 \boxed{\begin{array}{l}\alpha^y \\ \text{UNDERGRAD}(y)\end{array}} >} \quad \{2 < 1\}$$

Since only one path is associated to the scope forest in (19), we can derive from it the corresponding interpretation, as shown below.

4.3 Negation and Indefinites

Before introducing the rules of inference I need to fill in a few details. The first question is what to do with other operators with scope, like negation. The answer is that the tools introduced so far are in fact sufficient to represent scope ambiguities originated by negation, provided that we also index the negation operator. The representation of the sentence *John doesn't have a car*, for example, will be the scope forest (20).[6]

(20)
$$\varphi \boxed{\text{j} \ \neg_1\text{HAVE} < \mathbf{a}_2 \boxed{\begin{array}{l}\alpha^x \\ \text{CAR}(x)\end{array}} >} \quad \emptyset$$

The second question is how to make the 'path' idea work with sentences like (1), since the representation for indefinites in DRT_0 does not consist of a restriction and a scope. My answer is that it is possible to represent indefinite NPs with structures similar to those used for quantifiers - for example, it is possible to represent the disambiguated reading of (1) in which *every student* scopes over *an undergrad* as in (21) - without changing the properties that indefinites have in standard DRT.

(21)
$$\boxed{\boxed{\begin{array}{l}\alpha^x \\ \text{MALE-STUDENT}(x)\end{array}} \xrightarrow{\mathbf{every}_1} \boxed{\boxed{\begin{array}{l}\alpha^y \\ \text{UNDERGRAD}(y)\end{array}} \xrightarrow{\mathbf{a}_2} \boxed{\text{DATES}(x, y)}}}$$

Two such properties seem especially important. The first property is that indefinites, unlike quantifiers, are not subject to the scope constraint, as clearly shown by the contrast in acceptability between (22a) and (22b):

(22)
a. A dog_i came in. It_i sat under the table
b. Every dog_i came in. ??It_i sat under the table

The work on relational semantics and Dynamic Predicate Logic has however shown that we can model this distinction semantically, and still represent indefinites as having a scope [Rooth, 1987; Schubert and Pelletier, 1988]. It is enough to introduce two separate classes of DRS's: one of **referential** DRSs[7], used to represent determiners like *a*, *the*, and the pronouns; another of **quantified** DRSs, used to represent determiners like **most** and **every**. Both classes of DRS's will have

[6] This method can also be extended to modal operators.

[7] Bad as this name sounds, the only alternative that came to my mind was 'article DRS's', which is also misleading, because pronouns will also be represented in this way.

a restriction and a scope, but they will have different semantic properties: in particular, referential NPs will have the same accessibility properties that they have in DRT_1.

A second reason for not treating indefinites as quantifiers is that unselective operators like the universal and the conditional seem able to bind indefinites, but not generalized quantifiers. Again, this does not prevent a representation in which indefinites have a restriction and a scope. If desired[8], one can achieve the same semantic effects of DRT by giving to generalized quantifiers the capability of imposing constraints on the set of satisfying bindings, as shown in more detail in section 6.

In a word, representing indefinite NPs as in (21) won't require a change in the way they participate to anaphoric relations, nor will require a change in the properties of generalized quantifiers, unless this change is otherwise motivated. If desired, we could even define 'simplification rules' for transforming the structures associated to referential NPs into the representation more traditionally associated to indefinites in DRT.

A note on the notation: in order to distinguish between referential NPs and quantifiers in the scope forest representation, I will use parentheses instead of angle brackets for referential NPs; brackets will be used when any determiner is possible. The scope forest into which (1) is translated will be as follows:

(23) $$\boxed{<\mathbf{every}_1\ \boxed{\begin{array}{l}\alpha^x\\ \mathrm{MALE\text{-}STUDENT}(x)\end{array}}\ >\ \mathrm{DATES}\ (\ \mathbf{a}_2\ \boxed{\begin{array}{l}\alpha^y\\ \mathrm{UNDERGRAD}(y)\end{array}}\ )}$$

4.4 Plural Anaphora To Quantifiers

While intersentential singular anaphora to every-NPs and indefinites in the scope of a quantifier is subject to a number of restrictions[9], intersential plural anaphora is generally possible, as shown by the contrast between (22b) and (24), as well as by the contrast between (25a) and (25b).

(24) Every dog_i came in. They_i sat under the table.

(25) a. Every person with a dog_i came in. ??It_i was put under the table.
b. Every person with a dog_i came in. They_i were put under the table.

Without attempting a complete description of the phenomenon, I'll incorporate into the inference rules some facts about plural anaphora to quantifiers which will be useful to disambiguate on the basis of referential information.

I'll borrow the necessary notation from Link's **LP** logic [Link, 1987]. Link proposes a model in which the universe of discourse is not a set, but a complete semilattice $\langle E, \bigvee\rangle$ which contains all the 'sums' of the (atomic) individuals of a set $A \subset E$. With this model we can give an interpretation to the logical predicates $\mathrm{ATOM}(x)$, true iff the value associated to x in Link's model is an element of A, and $\mathrm{GROUP}(x)$, true iff that value is in $E - A$.[10] The following lemma holds:

Lemma 4.1 *For every marker x it is either the case that $\mathrm{ATOM}(x)$ or that $\mathrm{GROUP}(x)$, but not both.*

[8] That is, leaving aside the well-known objections raised against the unselective quantification account: the *proportion problem*, presented in more detail in section 6, puts under discussion the claim for generalized quantifiers, while Schubert and Pelletier [Schubert and Pelletier, 1988] present counterexamples to the claim for conditionals.

[9] Roberts discusses some cases in which it is possible in [Roberts, 1987]; see also [Poesio and Zucchi, 1991].

[10] Link's model has already been proposed by Kamp as a way to represent conjoined and plural NPs in DRT [Kamp, 1988]. Both ATOM and GROUP are mine.

The semantics of quantified DRS's like $K_1 \overset{\text{every}_i}{\rightarrow} K_2$ will be defined in terms of a **distancing** operation to be performed on the set of pairs $\langle f, g\rangle$ of embeddings such that f verifies the truth conditions of $K_1 \overset{\text{every}_i}{\rightarrow} K_2$ and g is one of the extensions of f which associate values to the markers which verify both restriction and scope. If we think of embeddings as ways of encoding situations, this operation is like a 'change in perspective': in the situation resulting from distancing we do not perceive the individual events and the single objects any more, but only the situation in its totality and the sets of objects involved. This means that after distancing, only the **projections** of the NPs, that is, the sets of objects playing certain roles in the global situation, are available for discourse anaphora. The contrast between (22a) and (22b), as well as the acceptability of (24), according to this account, are due to the fact that the projection of *a dog* in (22a) is a unique individual, and therefore available for individual anaphora, while the projection of *every dog* in (22b) and in (24) is the set of all dogs[11].

4.5 Inference Rules for Scope Disambiguation

At this point, I can begin to answer the questions: How can we infer the intended scope relations? Do we really need to infer them completely? What consequences can we infer from a non-disambiguated interpretation? The rules presented in this section are a way for using disambiguation information to infer the intended scope relations. The information about scope relations comes from a variety of sources. Three kinds of sources seem especially important:

1. Logical facts, like the fact that the sentence *A male student is dating an undergrad* has only one intepretation.

2. Anaphoric facts: If sentence (1) is followed by *They meet them at parties*, and we may conclude that either *them* or *they* is anaphoric to *an undergrad* in (1), that is, we may conclude that the projection of *an undergrad* in (1) is not a single person, but a group of people, then we may also conclude that *every male graduate* scopes over *an undergrad.*

3. World knowledge. For example, one may use facts about the social rules of dating to infer that the most likely interpretation of (1) is the one in which *every student* takes scope over *an undergrad.* As this very example shows, however, most of this information cannot be taken as conclusive, and therefore rules of this type are only appropriate with a logic which allows for revisions. My ideas on this problem are still tentative, and will be discussed at the end of the paper.

I will present an example of scope forest reduction rule based on 'logical' facts and two examples of rules based on 'anaphoric' facts. There is no pretense of completeness: the only reason why I give these specific rules is that they will get the examples in section 5 through. I'll then present the **Scope Forest Elimination** rule. The rules will be in the format specified in section 3.

[11] This account provides a justification for the operation of *summation* introduced in [Kamp, 1988], but of course doesn't solve the well-known problems raised by (26a) and (26b).

(26) a. Each student walked to the stage. He shook hands with the dean and left. (Partee)
b. Each Italian loves his car. He rides it every Sunday.

In [Poesio and Zucchi, 1991] we propose that distancing is blocked if the discourse has a certain structure - typically, the sentence which contains the anaphoric reference is an elaboration of a generic description or the continuation of an episode along a known 'script'.

ROR (Referential Over Referential) : This rule reflects the logical fact that referential NPs do not create scope ambiguities: in *A man saw a dog*, for example, the relative scope of *A man* and *a dog* does not matter.

$$\frac{\varphi\ \boxed{(\ \mathbf{d}_i\ \boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}}\ )\ \ \mathrm{R}\ \ (\ \mathbf{d'}_j\ \boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}}\ )}\qquad OC}{\varphi\ \boxed{(\ \mathbf{d}_i\ \boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}}\ )\ \ \mathrm{R}\ \ (\ \mathbf{d'}_j\ \boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}}\ )}\qquad OC \cup \{i < j\}}$$

RAOQ (Referential Atom Over Quantifier) : This rule allows the reduction of the scope forest associated with sentences like *Every male student dates an undergrad* once it has been concluded that *an undergrad* refers to a single individual, that is, the projection of *an undergrad* is an atom in Link's sense. It is worth remembering that because of the way the denotation of the scope forest is defined, y gets different values 'inside' and 'outside' of a quantified DRS (therefore, of a scope forest): 'outside' it denotes the *projection* of the referential NP.

$$\frac{\begin{array}{l}\varphi\ \boxed{<\ \mathbf{d}_i\ \boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}}\ >\ \ \mathrm{R}\ \ (\ \mathbf{d'}_j\ \boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}}\ )}\qquad OC\\ \mathrm{ATOM}(y)\end{array}}{\varphi\ \boxed{<\ \mathbf{d}_i\ \boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}}\ >\ \ \mathrm{R}\ \ (\ \mathbf{d'}_j\ \boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}}\ )}\qquad OC \cup \{j < i\}}$$

QORG (Quantifier Over Referential Group) : This rule enables us to conclude, from the fact that the projection of an indefinite NP is a group and the indefinite NP is in a scope forest with a quantifier, that the quantifier takes wide scope. (Consider for example the case when (1) is followed by *They meet them at parties.*)

$$\frac{\begin{array}{l}\varphi\ \boxed{<\ \mathbf{D}_i\ \boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}}\ >\ \ \mathrm{R}\ \ (\ \mathbf{D'}_j\ \boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}}\ )}\qquad OC\\ \mathrm{GROUP}(y)\end{array}}{\varphi\ \boxed{<\ \mathbf{D}_i\ \boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}}\ >\ \ \mathrm{R}\ \ (\ \mathbf{D'}_j\ \boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}}\ )}\qquad OC \cup \{i < j\}}$$

Last but not least, we need to be able to derive a disambiguated DRS from a scope forest associated to a single path. The simplest way for doing this is to introduce a rule of inference whose premise is a scope forest associated to a single path, as follows:

SFE (Scope Forest Elimination) : a scope forest which associated to a single path can be replaced by the corresponding interpretation.

$$\frac{\varphi\ \boxed{<\mathbf{d}_i\ \boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}}\ >\quad \mathrm{R}\quad <\mathbf{d'}_j\ \boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}}\ >}\qquad \{i<j\}}{\boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}}\ \xrightarrow{\mathbf{d}_i}\ \boxed{\boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}}\ \xrightarrow{\mathbf{d'}_j}\ \boxed{\mathrm{R}(i,j)}}}$$

With this way of writing the rules it will be simple to show how the derivations work in section 5, but, of course, one would then need one such rule for each permutation of the indices - 120 rules for a sentence with 5 quantifiers, for example. Plus, one would require one such rule for every number of arguments. In a word, it would seem that the combinatorial explosion that I was throwing out of the door is coming back through the window.

It is not so in practice, however. First of all, the particular notation for scope forests I have been using has been chosen to preserve the similarity with the logical forms proposed in the NLP literature as much as possible. With this notation, however, one can only write rules which apply to scope forests with a fixed number of arguments, and these rules are normally asymmetrical, in the sense that the argument position is significant. It should be clear however how rules for n-arguments scope forests could be written, as well as rules in which the argument position is reversed. There are notations in which more general rules can be written, but I preferred not to use them, since they are pretty opaque.

As for the potentially more dangerous problem of requiring $n!$ rules to disambiguate a scope forest with n arguments, a moment's thought will reveal that all that is really needed to trigger these rules is a way of checking that the ordering constraints define a total order, that is, there is a path of length $n-1$ from an index to another index. The inference procedure can do that without really going through $n!$ rules. I have a very simple minded algorithm which can add a new constraint and discover if the constraints define a total order in $O(n^2)$. Once it has been determined that the constraints define a total order, the operators can be 'extracted' from the scope forest one after the other in linear time using a procedure similar to the one that will be used in section 6 to define the set of paths associated to a scope forest.

4.6 Reasoning with Scope Forests

I mentioned before three reasons for assigning a semantics to unscoped logical forms: first, one doesn't need to compute all the possible interpretations; second, one can use information which comes later in the discourse to disambiguate; third, if one is able to draw the interesting conclusions on the basis of the ambiguous sentence alone, one may not need to disambiguate at all.

So far, however, I haven't said anything about how to go about performing inferences not related to disambiguation. We need to define inference rules analogous to first order logic's Universal Instantiation (UI) and Existential Generalization. It's easy to see how such rules can be defined (and semantically justified) in the framework I have been proposing. I will give as an example the scope forest version of UI; 'weak' versions of Existential Generalization and Existential Instantiation can be defined in the same way.

WUI (Weak Universal Instantiation) : from the sentences *Every male student dates an undergrad* and *John is a male student* conclude *John dates an undergrad.*

$$\dfrac{\begin{array}{l}\varphi\ \boxed{<\mathbf{every}_k\ \boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}}\ >\ \mathrm{R}\ \ (\ \mathbf{a}_l\ \boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}}\ )}\\ \mathrm{P}(b)\end{array}}{\boxed{\boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}}\ \xrightarrow{\mathbf{a}_l}\ \boxed{\mathrm{R}(b,y)}}}\ \ \mathrm{OC}$$

Also, similar rules old when a is replaced by *many*, *most*, etc.[12]

5 Reasoning with an Ambiguous Logic: Examples

In this section I will show how one can formally derive an unambiguous interpretation from a scope forest using the inference rules presented in the previous section, as well as facts about world knowledge and anaphoric relations. I will not give explicit examples of application of WUI, for reasons of space; hopefully, the derivations presented below will be explicit enough that the reader will be able to reconstruct the derivation of, e.g., *John dates and undergrad* from the scope forest representation of *Every male student dates an undergrad* and of *John is a male student.* I will instead take advantage of the examples to show how scope disambiguation can interact with other discourse disambiguation processes, namely, reference disambiguation. I will first present a simple model of reference disambiguation, then show the derivations.

5.1 An Elementary Model of Reference Disambiguation

The set of possible anaphoric antecedents of a pronoun x (its **anchoring set**) initially includes all the markers accessible to x according to a definition of accessibility which is essentially that of DRT, modulo the accessibility of plural markers for quantifiers. The initial anchoring set does not include those markers ruled out by binding constraints [Reinhart, 1983], since the parser introduces a disjointness condition $x \neq y$ for each such marker y.

I will need a logical predicate for talking about accessibility in the object language. The relation between two markers $x_i \prec x_j$ holds whenever x_i is accessible from x_j, that is, whenever x_i is introduced 'before' x_j. Semantically, $x_i \prec x_j$ will be defined to hold whenever $i < j$ (remember that each new marker has a greater index than any of the markers introduced before, and that no two markers are allowed to have the same index). $x_i \prec x_j$ is reflexive and transitive but not symmetrical. I will also make the simplifying assumption that if the marker x is the representation of a pronoun, then x is coindexed with one of the accessible markers. This assumption is encoded by the following axiom[13]:

[12] This was pointed out to me by Len Schubert.

[13] This axiom is too strong in general, but will do for my purposes. Consider the sentence *He came in*, and suppose that John and Bill are the only available referents. Using the disjunction method, we would obtain as a representation of the sentence that it is either the case that John came in or that Bill came in. Imagine, however, that neither John nor Bill came in. The theory would then predict that *He came in* is false, which doesn't seem the right prediction: most people would conclude instead that the referent of *he* is neither John nor Bill. This is, I think, yet another

Axiom 5.1 *Let $x_1 \ldots x_n$ be all the markers introduced before x_{n+1} (that is, all the markers for which $x_j \prec x_{n+1}$ is true), and let x_{n+1} be introduced by a pronoun. Then $x_{n+1} = x_1 \vee \ldots \vee x_{n+1} = x_n$ is also true.*

The first reference disambiguation rule adds new disjointness conditions, thus eliminating elements from the anchoring set. Let the * operator be defined as follows:

(*OP) $\textsc{p}^*(x) \equiv_{\text{def}} (\textsc{atom}(x) \wedge \textsc{p}(x)) \vee (\textsc{group}(x) \wedge \boxed{\begin{array}{l}\alpha^y\\ y \in x\end{array}} \xrightarrow{\textbf{every}_i} \boxed{\textsc{p}(y)}\ )$

The ASTR rule says that if two markers x and y are not of the same type, that is, a predicate $\textsc{p}$ is true of the marker x but not of y, then y and x are disjoint.

ASTR (Type Reduction) :

$$\frac{\textsc{p}^*(x),\ \ \neg\textsc{p}^*(y)}{x \neq y}$$

When the anchoring set of x reduces to just one element z the **parameter anchoring** (PA) rule applies. Parameter anchoring lets us infer a coindexing relation $x = y$ between a marker x and a marker y whenever every other marker z accessible from x has been found to be distinct from x:

PA (Parameter Anchoring) :

$$\frac{\begin{array}{l} y \prec x \\ \boxed{\begin{array}{l}\alpha^z\\ z \prec x\\ y \neq z\end{array}} \xrightarrow{\textbf{every}_i} \boxed{x \neq z} \end{array}}{x = y}$$

5.2 Disambiguation By Deduction

In this section I demonstrate the ideas presented up to this point by showing how each disambiguated interpretation of sentence (1) can be derived from the scope forest representation, given the appropriate context. The following theorem (an simple corollary of a lemma presented in section 4) will be used in the derivations:

Theorem 5.1 *Given any two markers x and z, if $\textsc{atom}(z)$ is true and $x = z$ is true, then $\textsc{atom}(x)$ is also true.*

argument against using identifying ambiguity with disjunction.

```
1. [<every 1 [x male-student(x)]> dates (a 2 [y undergrad(y)])], 0
                     (translation of ‘Every male student dates an undergrad’)

2. x <> y            (translation of ‘Every male student dates an undergrad’,
                      possibly incorporating some form of indefiniteness effect)

3. [speaker met (her 3 [z woman(z)])], 0
                     (translation of ‘I met her’)

4. x < z

5. y < z

6. woman*(z)         (part of the translation of ‘I met her’)

7. ~woman*(x)        (world knowledge)

8. z <> x            (4, 6, 7, ASTR)

9. z = y             (5, 2, 4, 8, PA)

10.atom(z)           (world knowledge)

11.atom(y)           (9, 10, Theorem 5.1)

12.[<every 1 [x male-student(x)]> dates (a 2 [y undergrad(y)])], {2 < 1}
                     (1, 11, RAOQ)

13.[y undergrad(y)] -a-> [[x male-student(x)] -every-> [x dates y]]
                     (12, SFE)
```

Figure 1: Wide scope for *an undergrad*

Let us consider again sentence (1), repeated here for convenience.

(28) Every male student dates an undergrad.

Let us now suppose that sentence (28) is followed by the sentence in (29).

(29) I met her.

In Fig. 1 I show how one can deduce a wide scope reading for *an undergrad* in (28) using the fact that it is coindexed with *her* in (29). Next, an example of disambiguation in which *every male student* takes wide scope. Let us now suppose that (28) is followed by the following sentence:

(30) They meet them at parties.

The derivation of a wide scope reading for *every male student* is shown in Fig. 2.

```
1. and 2.           as before.

3. [<they 3 [z group(z)]> meet-at-parties <they 4 [w group(w)]>], 0
                    (translation of 'They meet them at parties.')

4. and 5.:          as before

6. z <> w           (translation of 'They meet them at parties.')

7. w = x or w = y   (axiom 5.1)

8. z = x or z = y   (axiom 5.1)

9. y = w or y = z   (6, 7, 8)

10.group(z)         (part of the meaning of 'they')

11.group(w)         (part of the meaning of 'they'

12.group(y)         (9, 10, 11)

13.[<every 1 [x male-student(x)]> dates (a 2 [y undergrad(y)])], {1 < 2}
                    (1, 12, QORG)

14.[x male-student(x)] -every-> [[y undergrad(y)] -a-> [x dates y]]
                    (13, SFE)
```

Figure 2: Narrow scope for *an undergrad*

6 A More Formal Presentation of DRT$_1$

6.1 Syntax

The main syntactic differences between DRT$_0$ and DRT$_1$ are the following:

1. Universally quantified sentences and conditionals are now represented as two different classes of expressions: the class of **quantifier** DRSs, and the class of **connective** DRSs.

2. A new class of complex conditions is introduced, the class of **referential** DRSs, to represent indefinites, definites and pronouns. These kinds of NPs are therefore syntactically separated from proper names.

3. Another new kind of complex condition is introduced, the **scope forest**.

4. Operators like quantifiers and negation are given **indices**.

The set of symbols of DRT$_1$ includes, in addition to the set of symbols of DRT$_0$, a set of **indices** $0 \ldots i \ldots$; a set **QDet** = {**every**, **most** } of **quantifier operators**; and a set **RDet** = {**a**, **the**, **he**, **it**, **she**, **they**} of **referential operators**. The set of expressions of DRT$_1$ consists of marker introducers (defined as before), conditions, and DRSs. In addition to the unary, binary and coindexing conditions of DRT$_0$, DRT$_1$ includes the following types of conditions:

1. **disjointness conditions** of the form $x \neq y$, where x and y are markers.

2. **accessibility conditions** of the form $x \prec y$, where x and y are markers.

3. the **structural conditions** $\mathrm{ATOM}(x)$ and $\mathrm{GROUP}(x)$, where x is a marker.

4. **negated** DRSs, that are expressions of form $\neg_i K$, where K is a DRS and i an index;

5. **connective** DRSs, like $K_1 \stackrel{\textbf{if-then}_i}{\rightarrow} K_2$ (used to represent the conditional), where K_1 and K_2 are DRSs;

6. **quantifier** DRSs, which are expressions of the form

$$\boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}} \xrightarrow{\mathbf{d}_i} \boxed{\mathrm{Q}(x)}$$

   where $\mathbf{d} \in \mathbf{QDet}$, and K_1 and K_2 are DRSs; K_1 will be called **restriction**, K_2 **scope**, and x the **main marker**;

7. **referential** DRSs, again of the form $K_1 \stackrel{\mathrm{d}_i}{\rightarrow} K_2$, where $\mathbf{d} \in \mathbf{RDet}$, and K_1 and K_2 are DRSs; again, K_1 will be called restriction and K_2 will be called scope;

8. **scope forests**, which are expressions of one of the following types:

$$\varphi \boxed{< \mathbf{d}_i \boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}} > \quad \mathrm{R} \quad < \mathbf{d}'_j \boxed{\begin{array}{l}\alpha^y\\ \mathrm{Q}(y)\end{array}} >}\ \mathrm{OC}$$

$$\varphi \boxed{< \mathbf{d}_i \boxed{\begin{array}{l}\alpha^x\\ \mathrm{P}(x)\end{array}} > \quad \mathrm{R}}\ \mathrm{OC}$$

   where $\mathbf{d}$ and $\mathbf{d}'$ are in $\mathbf{QDet} \cup \mathbf{RDet}$; R is either a relation or a *negated relation* of the form $\neg_i \mathrm{R}'$ where R' is a relation; and *OC* is a set $\{i \prec j, \ldots k \prec l\}$ of **ordering constraints** among the indices of the operators in the scope forest.

A **formula** of DRT_1 is a DRS with no free markers, and in which no two operators are given the same index. The donkey sentence *Every farmer who owns a donkey beats it* is represented as in (31) (which the reader should compare to the DRT_0 representation (14), section 3).

$$(31)\quad \boxed{\boxed{\begin{array}{l}\alpha^x\\ \mathrm{FARMER}(x)\\ \boxed{\begin{array}{l}\alpha^y\\ \mathrm{DONKEY}(y)\end{array}} \xrightarrow{\mathbf{a}_2} \boxed{\mathrm{OWNS}(x,y)}\end{array}} \xrightarrow{\mathbf{every}_1} \boxed{\boxed{\begin{array}{l}\alpha^z\\ z = y\end{array}} \xrightarrow{\mathbf{it}_3} \boxed{\mathrm{BEATS}(x,z)}}}$$

6.2 Semantics

A model M for DRT_1 is a pair $\langle U, F\rangle$: F is an interpretation function, and U is a Link-type boolean algebra $< D, A, \leq, \sqcup >$ in which $\leq$ models the inclusion relation, A is the set of *atoms* of D, i.e., those elements of D such that no other element is included in them, and $\sqcup$ models sum ($x \sqcup y$ is the minimal element of U which includes both x and y).

The denotation of a DRT_1 expression with respect to M is defined as follows. The denotation of the basic conditions and of the marker introducers is the same as in DRT_0. The new atomic conditions have the following denotations:

1. $\|x \neq y\|^M = \{\langle f, f\rangle |\ f(x) \neq f(y)\}$
2. $\|\text{ATOM}(x)\|^M = \{\langle f, f\rangle |\ f(x) \in A\}$
3. $\|\text{GROUP}(x)\|^M = \{\langle f, f\rangle |\ f(x) \in U - A\}$
4. $\|x_i \prec x_j\|^M = \emptyset$ if $i > j$, $\{\langle f, f\rangle |\ x_i, x_j \in \text{DOM}(f)\}$ otherwise.

The denotation of the basic DRS changes, since accessibility is now controlled by discourse structure and the type of NP. Rather than pairs $\langle f, f\rangle$, a basic DRS will denote pairs $\langle f, g\rangle$ where g contains additional values for the markers introduced by the DRS.

1. $$\left\| \boxed{\begin{array}{l} \alpha^{x_1}, \ldots, \alpha^{x_n} \\ C_1 \\ \vdots \\ C_m \end{array}} \right\|^M = \{\langle f, g\rangle |\quad f \subset f_1 \subset \ldots \subset f_n \subset g_1 \subseteq \ldots \subseteq g_{m-1} \subseteq g, \\ \langle f, f_1\rangle \in \|\alpha^{x_1}\|^M \text{ and } \ldots \langle f_{n-1}, f_n\rangle \in \|\alpha^{x_n}\|^M \text{ and} \\ \langle f_n, g_1\rangle \in \|C_1\|^M, \ldots \langle g_{m-1}, g\rangle \|C_m\|^M\}$$

The semantics of quantifier and referential DRSs will be a straightforward extension of the semantics of the conditional DRS in DRT_0. Let us first introduce the following definitions. Let $K_1 \overset{\mathbf{d}_i}{\rightarrow} K_2$ be a quantifier or referential DRS, with determiner **d**, index i, and main marker x. Let f be an embedding such that $x \notin \text{DOM}(f)$. The set of DRS-*Satisfying Embeddings*, DSE, is defined as follows:

$$\text{DSE}(f, K) = \{h |\ \langle f, h\rangle \in \|K\|^M\}$$

and extended as follows to sets of embeddings:

$$\text{DSE}(\{f_1, \ldots, f_n\}, K) = \{h |\ \langle f_i, h\rangle \in \|K\|^M \text{ for some } i \in 1 \ldots n\}$$

I need now to make the notion of 'distancing' introduced in section 4 more precise. As said there, distancing has the effect of 'changing the perspective' on a certain situation, that is, making available 'outside' the DRS only the sets of participants to a situation 'as a whole', so that whereas before the closure the markers are available for singular anaphora, after they are only available for plural anaphora.

Definition 6.1 *Let E be a set of embedding pairs $\{\langle f, g\rangle\}$. $\text{DISTANCING}(E)$ is the set of embedding pairs $E' = \{\langle f, h\rangle\}$ defined as follows. Let $x_i \ldots x_j$ be the markers over which g extends f. Then $\langle f, h\rangle$ will be in $\text{DISTANCING}(E)$, where h is the embedding which is like g for the markers up to x_{i-1}, and then for $l = i \ldots j$, $h(l) = \sqcup\ h'(l)$, for any h' s.t. $\langle f, h'\rangle \in E$.*

For example, an output embedding h in the denotation of a quantified DRS like

$$\boxed{\begin{array}{l} \alpha^x \\ \text{P}(x) \end{array}} \overset{\mathbf{every}_i}{\longrightarrow} \boxed{\text{Q}(x)}$$

gives as value to the marker x the group of all individuals which satisfy both the restriction and the scope of the quantified DRS.[14] We can now specify the denotations of quantified and referential DRSs:

1. $\|K_1 \stackrel{\text{every}_i}{\rightarrow} K_2\|^M = \text{DISTANCING}(\{ \langle f, g\rangle$, where f is an embedding such that for every $h \in \text{DSE}(f, K_1)$ there is an h' s.t. $h' \in \text{DSE}(h, K_2)$, and $g \in \text{DSE}(\text{DSE}(f, K_1), K_2)\})$

2. $\|K_1 \stackrel{\text{most}_i}{\rightarrow} K_2 \|^M = \text{DISTANCING}(\{ \langle f, g\rangle|$ where f is an embedding such that for most $h \in \text{DSE}(f, K_1)$ there is an h' s.t. $h' \in \text{DSE}(h, K_2)$, and $g \in \text{DSE}(\text{DSE}(f, K_1), K_2)\})$[15]

3. $\|K_1 \stackrel{\text{a}_i}{\rightarrow} K_2\|^M = \{ \langle f, g\rangle|\ \text{DSE}(\text{DSE}(f, K_1), K_2) \neq \emptyset$ and $g \in\ \text{DSE}(\text{DSE}(f, K_1), K_2)\}$

4. $\|K_1 \stackrel{\text{he}_i}{\rightarrow} K_2\|^M = \{ \langle f, g\rangle|\ f(x) = f(y)$ for some marker y and $\langle f, g\rangle \in \|K_2\|^M\}$

(Analogous rules could be developed for $K_1 \stackrel{\text{she}_i}{\rightarrow} K_2$ and $K_1 \stackrel{\text{they}_i}{\rightarrow} K_2$, of course.) Next, the denotations of the negated DRS and the conditional, almost identical to those in DRT_0::

1. $\|\neg_i K\|^M = \{\langle f, f\rangle|$ there is no g such that $\langle f, g\rangle \in \|K\|^M\}$

2. $\|\ K_1 \stackrel{\text{if-then}_i}{\rightarrow} K_2\|^M = \{\langle f, f\rangle|$ for every h s.t. $\langle f, h\rangle \in \|K_1\|^M$ there is an l s.t. $\langle h, l\rangle \in \|K_2\|^M$, and g is one such $l\}$

Finally, in order to define the semantics of scope forests, we have to specify more precisely the set of paths associated to them. This definition will be given in terms of an *extraction procedure*.

[14] This semantics is correct for the examples in the previous section, but also predicts that in texts like *Every man who has a donkey beats it. They hate him.*, *They* would refer to the set of all donkeys beaten by any man. To fix this problem it would be necessary to stipulate an ambiguity.

[15] This definition suffers from the so called *proportion problem* [Heim, 1982; Rooth, 1987]. Let us consider a model with 100 farmers, 99 of which own a donkey and don't thrive whereas one, Pedro, owns 1000 donkeys and thrives. The definition of the semantics of $K_1 \stackrel{\text{most}_i}{\rightarrow} K_2$ above would predict that this model would make *Most farmers who own a donkey thrive* true, which is rather counterintuitive. The reason is that if the restriction is a DRS K like

$$\boxed{\begin{array}{l} \alpha^x \\ \text{FARMER}(x) \\ \boxed{\begin{array}{l} \alpha^y \\ \text{DONKEY}(y) \end{array}} \stackrel{\mathbf{a}_2}{\longrightarrow} \boxed{\text{OWNS}(x, y)} \end{array}}$$

$\text{DSE}(f,K)$ will include a different embedding for each *pair* of values for the markers x and y, and the fact that Pedro owns 1000 donkeys will make him a participant in 1000 pairs. The source of the problem, in short, is that the pairs are counted, rather than the number of possible distinct values that may be given to x. R. Root presented a solution to the problem which could easily be accomodated in my framework. Her idea was to partition a set of embeddings into equivalence classes according to the value that the embeddings associate to x:

> $\text{EP}(E, x)$, where E is a set of embeddings, and x is a marker, $= \{ E_a|$ each $h \in E_a$ is a member of E, and for all $h \in E_a\ h(x) = a$ for some $a \in U\}$

and then replacing the definition of $K_1 \stackrel{\text{most}_i}{\rightarrow} K_2$ above with the one below:

- $\|K_1 \stackrel{\text{most}_i}{\rightarrow} K_2 \|^M = \{ \langle f, g\rangle|$ for most $E \in \text{EP}(\text{DSE}(f, K_1), x)$ there is an $h \in E$ s.t. $h' \in \text{DSE}(h, K_2)$ for some h' $\ldots\}$ (g is defined as before)

However, even if this definition does not suffer from the problem above, it does not explain why the semantics of *most* should be defined in terms of equivalence classes, while the semantics of *every* should be based on pairs. A more illuminating solution would be to assume an ambiguity, originated by the possibility of choosing different farmers-owning-donkeys 'cases' in the restriction of a quantifier; developing this solution in detail, however, would go outside the scope of this paper.

Definition 6.2 *An* **extraction procedure** *is an algorithm which 'pulls out' all the operators from a scope forest $\varphi[i_1 \ldots i_n]OC$, one after the other, and returns a* **disambiguated** DRS *(a* DRS *not a scope forest), under the constraint that if the ordering constraint $i < j$ is in OC, then the operator with index i is extracted before the operator with index j.*

What the 'extraction rules' do is fairly obvious and can be compared to what Quantifier Raising does. I will just give two examples:

negation:

$$\varphi \boxed{T_1 \neg_i \text{ R } T_2} \qquad \Rightarrow \neg_i \boxed{\varphi \boxed{T_1 \text{ R } T_2}}$$

indefinites:

$$\varphi \boxed{T_1 \text{ R } (\mathbf{a}_i \boxed{\begin{array}{l} \alpha^x \\ \text{P}(x) \end{array}})} \qquad \Rightarrow \boxed{\begin{array}{l} \alpha^x \\ \text{P}(x) \end{array}} \xrightarrow{\mathbf{a}_i} \boxed{T_1 \text{ R } x}$$

The first extraction rule extracts a negation operator from a scope forest, and produces as a result a negated DRS with the scope forest from which the operator has been extracted in the scope of the negation operator. T_1 and T_2 can be either unextracted terms or markers. The second rule extracts an indefinite term from a scope forest, and produces a referential DRS in whose scope there is a new scope forest in which the marker x has replaced the indefinite term. It is worth observing that nothing prevents having more than one extraction rule per determiner, or extraction rules which operate on the scope forest as a whole (this will take care of the additional readings of examples like *Two examiners marked six scripts*). An 'empty extraction rule' will replace a scope forest without operators left with a basic DRS.

Definition 6.3 *Let K be the scope forest $\varphi[i_1 \ldots i_n]OC$ with indices $i_1 \ldots i_n$ and ordering constraints OC. The* **paths associated to** *SF are all the disambiguated* DRS*'s K' which can be derived from K by an extraction procedure.*

We can now define the denotation of a scope forest as follows:

1. $\| K : \varphi[i_1 \ldots i_n]OC \|^M = \bigcup \|P\|^M$, for every path P associated to K

6.3 Inference Rules

The set of inference rules of DRT_1 includes, in addition to the traditional rules (Modus Ponens, Universal Instantiation, etc.) all the rules defined in section 4 and 5. The following theorem holds:

Theorem 6.1 *MP, UI, ROR, RAOQ, QORG, SFE, WUI, PA and ASTR are sound* DRT_1 *inference rules.*

7 Discussion

I have shown that relational semantics can be used to give a denotation to certain expressions, called scope forests, which are then used to represent scopally ambiguous sentences. As a consequence, certain types of scope disambiguation can be performed by a deductive procedure. I have also shown examples of such derivations. In this section I would like to discuss more in detail what the theory predicts and what it doesn't.

First of all, I do not claim that people do not disambiguate unless they can use world knowledge or referential information. This position is hard to defend: there seems to be psychological evidence that, at least for simple sentences like *Every boy climbed a tree*, people do have interpretation preferences [Kurtzman and MacDonald, 1991]. Giving a semantics to unscoped logical forms does not rule out scope forest reduction rules based on preferences, however, although these preferences should be incorporated as defeasible rules, with consequent problems in formalization.[16]

An interesting issue is whether these preferences always apply, that is, whether people always attempt to come out with one unambiguous interpretation. An interesting hypothesis is that the cost of disambiguating is a factor in determining whether or not people try to disambiguate. Once more, the theory just presented is neutral with respect to this question (although it would be most useful for the cases when disambiguating is very expensive). Kurtzman and MacDonald only studied sentences with two quantified NPs, and the only determiners that appeared in those sentences were *every*, *the*, *a* and *a different*. These quantifiers are 'easy to compute' - they can all be computed by DFAs [van Benthem, 1987]. It would be interesting to repeat the experiment with determiners which are 'harder' than the determiners above, such as *most* and *few*: for example, to try sentences like *Most male students date few undergrads.* Another source of complexity is the number of possible interpretations: the prediction would be that as the number of interpretations gets larger, it should be less likely that people disambiguate. This hypothesis could be tested by running experiments on sentences containing numerical quantifiers, such as *Two examiners marked six transcripts.*.

Another question left open by Kurtzman and MacDonald's experiments, is whether people, when they do disambiguate, compute only one interpretation (as predicted by the theory presented in this paper) or compute all interpretations in parallel, as predicted by the disjunction theory. The problem is that the timings of the experiments (approximately 3.5 seconds.) were not strict enough to guarantee that only one interpretation could be computed [Kurtzman, p.c.]. Testing people on sentences with numerical quantifiers would be an interesting test for this hypothesis also. The sheer number of interpretations should make it possible to verify whether indeed people compute them all and then 'filter', or rather compute a few (possibly using preference heuristics).

One should not read the theory as claiming that syntactic constraints play no role in determining scope relations either. As May [1985] argues, sentences like *Some student admires every professor* may be ambiguous in isolation, but in a VP-deletion context, like (32), the ambiguity evaporates.

(32) Some student admires every professor, but John doesn't.

All that my theory says is that scope determination is not an entirely syntactic operation. In this, it is similar to the theory proposed by Gawron and Peters in [Gawron and Peters, 1990](cfr. pages 53–56). In both cases, the determination of the scope of NPs is attributed to contextual factors (Gawron and Peters call them *circumstances*). The two theories are however trying to answer different questions. As I mentioned before, I am interested in questions like: How can we infer the intended scope relations? Do we really need to infer them completely? What consequences can we infer from a non-disambiguated interpretation? Instead, Gawron and Peters are concerned with questions like: How do SCOPES-OVER facts get into the circumstances? How do we capture the constraints on scoping and anaphora reflected, for example, by data about sloppy and strict anaphora?

[16] Kurtzman and MacDonald's results do seem to be in contrast with the predictions of approaches based on using the weakest possible interpretation, like Kempson and Cormack's, since people do not seem to favor the weakest interpretation, and actually at times have a definite preference for the strong one.

8 Conclusions

8.1 Status of the Implementation

A GPSG grammar for a fragment of English more or less equivalent to that of 'standard' DRT has been developed, which produces DRT_1 expressions. A parser which carries on the translation 'bottom up' in a fashion similar to that proposed by Johnson and Klein [1986] has also been implemented in PROLOG, together with a module which carries on the inferences presented in this paper, using a logical form very similar to that use in the Core Language Engine [Alshawi and van Eijck, 1989].

8.2 Future Work

A crucial question is, of course, whether we get any actual performance improvement by adopting this method, that is, whether inferencing with DRT_1 is so expensive as to offset the advantages gained from not having to produce all the interpretations. A general answer would require determining the complexity of reasoning with DRT_1, which hasn't been done yet. Even an experimental answer is hard to get, since that requires, in addition to implementing two or three different scope disambiguation methods (which we have actually already done in part), making sure that their coverage is actually the same, and elaborate appropriate tests.

A second, very important problem is to have some way of jumping to conclusions on the basis of uncertain knowledge, and of choosing between different possibilities. A logic of this form seems necessary to formalize the kind of preferences discussed by Kurtzman and MacDonald; in addition, we'll probably need it to represent the disambiguating information originated from lexical and world knowledge. For example, it is a lexical property of the predicate *comes with* that the argument in subject position takes scope over the argument in object position, so that *Every chess set comes with a spare pawn.* has no reading in which *a spare pawn* scopes over *Every chess set.* In order to express such 'defeasible disambiguation axioms' we have developed the **Disambiguation Schema**, whose general form is `S --> {i<j}`, where `S` is the scope of the quantifier with the strictest scope and `i` and `j` are indices of operators. For example, the disambiguation schema for the predicate *comes with* would be `[i comes-with j] --> {i < j}`. This schema is an abbreviation for a list of axiom schemata each of which adds the ordering constraint $i < j$ to a scope forest. Unfortunately, schemata of this kind do not always hold.

A question that needs to be addressed in more detail is the relation between scope ambiguities and, for example, collective/distributive ambiguities, in order to see better the relation between this approach and, for example, the approach proposed by Bunt. The reference interpretation model also needs to be extended. I'm in particular interested in studying sentences like *Every chess set comes with a spare pawn. It's taped to the bottom.*, in which the interface between reference and scope disambiguation seems totally different from the one presented in the examples in section 5.

Acknowledgments

I am grateful to my advisor, Len Schubert, Sandro Zucchi, James Allen, Jeff Pelletier and Graeme Hirst for their comments. I have profited considerably from discussions with Mats Rooth, Mark Gawron, John Nerbonne, David Israel, Dag Westerstahl, Fernando Pereira, Barbara Di Eugenio, Heinz Marburger, Michael Fliegner, and Rolf Eberle. I also thank an anonymous referee for editorial comments, suggestions to sharpen the focus of the paper, and for pointing out unclarities.

References

[Allen, 1987] Allen, J. 1987. *Natural Language Understanding.* Benjamin Cummings, Menlo Park, CA.

[Allen, 1991] Allen, J. 1991. Natural language, knowledge representation, and logical form. Technical Report 367, University of Rochester, Rochester, NY.

[Alshawi and van Eijck, 1989] Alshawi, H. and Eijck, J. van 1989. Logical forms in the Core Language Engine. In *Proc. ACL-89*, Vancouver, CA.

[Alshawi *et al.*, 1988] Alshawi, H.; Carter, D. M.; van Eijck, J.; Moore, R.C.; Moran, D. B.; F.C.Pereira, ; and Smith, A.G. 1988. Interim report on the SRI Core Language Engine. Technical Report CCSRC-5, SRI International, Cambridge, Cambridge, England.

[Barwise and Perry, 1983] Barwise, J. and Perry, R. 1983. *Situations and Attitudes.* The MIT Press.

[Barwise, 1987] Barwise, J. 1987. Noun phrases, generalized quantifiers and anaphora. In Gärdenfors [1987]. 1–30.

[Bunt, 1985] Bunt, H. 1985. *Mass Terms and Model Theoretic Semantics.* D. Reidel.

[Gärdenfors, 1987] Gärdenfors, P., editor 1987. *Generalized Quantifiers.* D. Reidel.

[Gawron and Peters, 1990] Gawron, J. M. and Peters, S. 1990. *Anaphora and Quantification in Situation Semantics*, volume 19 of *Lecture Notes.* CSLI.

[Groenendijk and Stokhof, 1990] Groenendijk, J.A.G. and Stokhof, M.J.B. 1990. Dynamic Montague Grammar. In Kálmán, L. and Pólos, L., editors 1990, *Papers from the 2nd Symposium on Logic and Language.* Akadémiái Kiadó, Budapest.

[Grosz *et al.*, 1987] Grosz, B.; Appelt, D.; Martin, P.; and Pereira, F. 1987. TEAM: An experiment in the design of transportable natural language interfaces. *Artificial Intelligence* 32(2):173–244.

[Heim, 1982] Heim, I. 1982. *The Semantics of Definite and Indefinite Noun Phrases.* Ph.D. Dissertation, University of Massachusetts at Amherst.

[Hobbs and Shieber, 1987] Hobbs, J. and Shieber, S. 1987. An algorithm for generating quantifier scopings. *Computational Linguistics* 13(1-2):47–63.

[Hobbs, 1982] Hobbs, J. 1982. Representing ambiguity. In *Proc. WCCFL-82.* 15–28.

[Hobbs, 1983] Hobbs, J. 1983. An improper treatment of quantification in ordinary English. In *Proc. ACL-83*, Cambridge, MA. 57–63.

[Hurum, 1988] Hurum, S. 1988. Handling scope ambiguities using domain-independent heuristics. Technical Report TR 88-12, University of Alberta, Edmonton, Canada.

[Johnson and Klein, 1986] Johnson, M. and Klein, E. 1986. Discourse, anaphora and parsing. Technical Report CSLI-86-43, CSLI, Stanford, CA.

[Kamp, 1981] Kamp, H. 1981. A theory of truth and semantic representation. In Groenendijk, J.; Janssen, T.; and Stokhof, M., editors 1981, *Formal Methods in the Study of Language.* Mathematical Centre, Amsterdam.

[Kamp, 1988] Kamp, H. 1988. Some remarks about the semantic representation of certain plural constructions of English, and suggestions for a procedural treatment of plural pronoun anaphora. Contribution to Acord Workshop.

[Kempson and Cormack, 1981] Kempson, R. and Cormack, A. 1981. Ambiguity and quantification. *Linguistics and Philosophy* 4(2):259–310.

[Kurtzman and MacDonald, 1991] Kurtzman, H. S. and MacDonald, M. C. 1991. Interpretation preferences for quantifier scope ambiguities. In *Proc. SALT-1.*

[Link, 1987] Link, G. 1987. Generalized quantifiers and plurals. In Gärdenfors [1987]. 151–180.

[May, 1985] May, R. 1985. *Logical Form in Natural Language.* The MIT Press.

[Partee, 1975] Partee, B. 1975. Comments on C.J.Fillmore and N.Chomsky's papers. In Austerlitz, D., editor 1975, *The scope of American linguistics.* De Ridder Press, The Netherlands.

[Poesio and Zucchi, 1991] Poesio, M. and Zucchi, S. 1991. On some alleged violations of the Scope Constraint. Submitted for Publication.

[Reinhart, 1983] Reinhart, T. 1983. *Anaphora and semantic interpretation.* Croom Helm, London.

[Roberts, 1987] Roberts, C. 1987. *Modal Subordination, Anaphora, and Distributivity.* Ph.D. Dissertation, University of Massachusetts at Amherst.

[Rooth, 1987] Rooth, M. 1987. Noun phrase interpretation in Montague Grammar, File Change Semantics, and Situation Semantics. In Gärdenfors [1987]. 237–268.

[Schubert and Pelletier, 1982] Schubert, L. K. and Pelletier, F. J. 1982. From English to Logic: Context-free computation of 'conventional' logical translations. *American Journal of Computational Linguistics* 10:165–176.

[Schubert and Pelletier, 1988] Schubert, L. K. and Pelletier, F. J. 1988. Generically speaking. In Chierchia, G. and Partee, B. H. and Turner, R., eds. *Properties, Types and Meaning, vol. II.* Kluwer, Dordrecht. 193–268.

[van Benthem, 1987] van Benthem, J. 1987. Towards a computational semantics. In Gärdenfors [1987]. 31–72.

[Webber, 1978] Webber, B. Lynn 1978. A formal approach to discourse anaphora. Report 3761, BBN, Cambridge, MA.