

Discourse Parallelism, Scope, and Ellipsis

Nicholas Asher

University of Texas at Austin

Daniel Hardt

Villanova University

Joan Busquets

IRIT/Université P. Sabatier

1 Introduction

It has frequently been observed that structural ambiguity does not multiply in contexts involving ellipsis:¹ that is, if there is an ambiguity associated with the antecedent of an ellipsis occurrence, that ambiguity must be resolved in the same way in both the antecedent and at the ellipsis site. The following example (Dalrymple et al., 1991) illustrates this with a case of quantifier scope ambiguity:

- (1) John gave every student a test, and Bill did too.

The antecedent clause allows two readings, depending upon whether *a test* or *every student* receives wide scope. However, if *a test* receives wide scope in the antecedent, it must also receive wide scope at the ellipsis site; conversely, if *every student* receives wide scope in the antecedent, it must receive wide scope at the ellipsis site as well.

Many accounts of ellipsis attempt to account for this observation in terms of the ellipsis recovery mechanism. However, similar effects are observed in cases that do not involve ellipsis, where the ellipsis recovery mechanism cannot apply, as in the following variant of (1):

- (2) John gave every student a test, and Bill gave every student a project.

To explain these facts, we will define a general parallelism constraint on related sentences in a discourse. This makes it possible to define an extremely simple recovery mechanism for ellipsis – an exact identity condition in which the recovered material and the antecedent are interpreted independently in their respective contexts, subject only to the general discourse constraints on parallelism. Furthermore, this approach makes it possible to give a uniform account for parallelism facts, whether or not there is ellipsis. In addition, our account explains the wide scope puzzle (Sag, 1976), illustrated by the following example:

- (3) A nurse saw every patient. Dr. Smith did too.

As observed by Sag, “a nurse” must take wide scope.

In what follows, we first consider the recovery mechanism for VP ellipsis. Next, we describe the basic parallelism constraints on related sentences in discourse, using the Segmented Discourse Representation Theory (SDRT) of (Asher, 1993). We show how these constraints capture well-known facts about scope parallelism in ellipsis, and we show how our account captures similar facts about parallel scope not involving ellipsis.² Next we show how our approach accounts for the wide scope puzzle. We then consider related work, and discuss certain extensions to our approach.

2 VP Ellipsis Recovery Condition

In our view, facts about parallel interpretation in ellipsis contexts reflect general constraints on related sentences in a discourse, rather than constraints on the specific mechanism for resolving ellipsis. This suggests that the ellipsis recovery mechanism is much simpler than is generally supposed. We propose that the recovery mechanism is a simple identity of structure, in which the recovered material is interpreted independently of the antecedent, subject only to general discourse parallelism conditions that apply to all related sentences.

This identity condition could be defined at a variety of levels of representation. The argument of this paper would be consistent with an identity condition defined at Surface Structure (SS), Logical Form (LF), or at the level of Discourse Representation Structure (DRS). For concreteness, we describe two possible identity conditions, at LF and DRS.

2.1 LF Recovery Condition

We assume that the LF representations are derived from SS representations by means of applications of a Quantifier Raising (QR) operation (as defined, for example, in (May, 1977)), which optionally adjoins quantified NP’s to S nodes. To implement our VPE recovery condition at LF, we propose an operation VP-copy, which can copy a VP to an empty VP node. We impose no ordering on QR and VP-copy³.

Consider the following example:

- (4) A nurse saw every patient. A doctor did too

[[A nurse] [past [see [every patient]]]] [[A doctor] did [] too].

The application of VP-copy results in

[[A nurse] [past [see [every patient]]]] [[A doctor] did [see [every patient]] too]

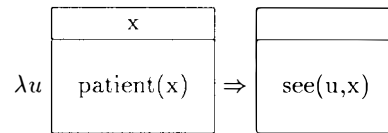
The application of QR in each clause results in one possible reading:

[[every patient] [[A nurse] [past [see [e]]]]]
 [[every patient] [[A doctor] [past [see [e]]]]]

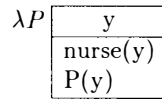
In fact, our recovery condition would permit four possible readings here. As we will see in section 4, the non-parallel scopings are ruled out by the general discourse parallelism constraint.

2.2 DRS Recovery Condition

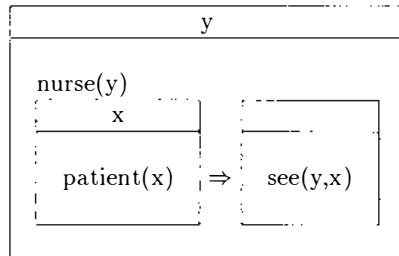
A similar effect can be obtained with a semantic recovery condition, in which VPE recovery and scope determination are freely ordered.⁴ We describe a recovery condition at the DRS level. To permit wide scope object readings in the DRS construction process, we permit type raising, to allow a certain flexibility in the order in which constituents are combined. We consider again example (4). The normal representation of the VP *saw every patient* would be:⁵



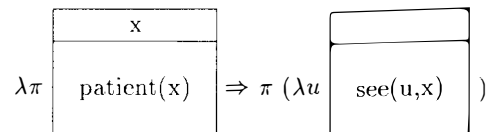
The subject *a nurse* is represented



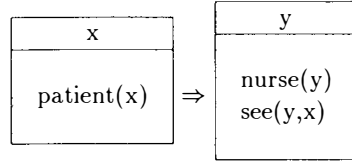
Application of the subject to the VP results in a narrow scope for the object:



An optional type raising rule gives the following representation for the VP:



Application of this VP representation to the subject results in a wide scope reading for the object:



Thus, we simply copy the VP representation constructed during the normal application of the DRS construction process. We allow a type raising operation as a way of obtaining wide scope object readings, and this type raising of VP representations can be interleaved with the VPE recovery operation.

For the purposes of this paper, either the LF or DRS recovery mechanism would be acceptable. We now turn to the basic discourse constraint that determines preferences among the various scopings permitted by our VPE recovery mechanism.

3 SDRT and the Maximization Constraint

Segmented Discourse Representation Theory (SDRT), (Asher, 1993; Lascarides and Asher, 1993) extends Kamp’s DRT (Kamp, 1980) by adding a more complex account of discourse structure. The constituents of discourse structure are *segmented DRSs* or SDRSs. These SDRSs are defined recursively out of DRSs and discourse relations, which are taken to be binary relations between propositions.

In resolving scope ambiguity, we claim that there is a preference to produce the *Maximal Common Theme (MCT)* of related discourse constituents. This is a modified version of the *Maximization Constraint* of (Asher, 1993).⁶ We now define Maximal Common Theme in terms of operations for eliminating or generalizing information in a DRS.

- *Operations*

We define four operations to eliminate or generalize information in a DRS:

1. Delete an atomic condition
2. Generalize an atomic condition (i.e., replace a relation R with a relation R’ where the denotation of R is a subset of the denotation of R’.)⁷
3. Systematically rename a bound discourse marker
4. Delete a discourse marker

We say that $K \rightsquigarrow K'$ iff K' is constructed from K by (0 or more) applications of rules 1-4.

- *Theme*: if $K \rightsquigarrow K'$, then we say that K' is a *theme* of K .
- *Maximal Common Theme (MCT)*: given two DRS's K, J the MCT is the DRS T such that $K \rightsquigarrow T$ and $J \rightsquigarrow T$, and for any other T' such that $K \rightsquigarrow T'$ and $J \rightsquigarrow T'$, $T \rightsquigarrow T'$.
- *Maximization Constraint*: in resolving a scope ambiguity within a pair of related discourse constituents, prefer the choice that produces the maximal MCT.

In the next section, we see how our constraint applies to the basic parallel scope examples.

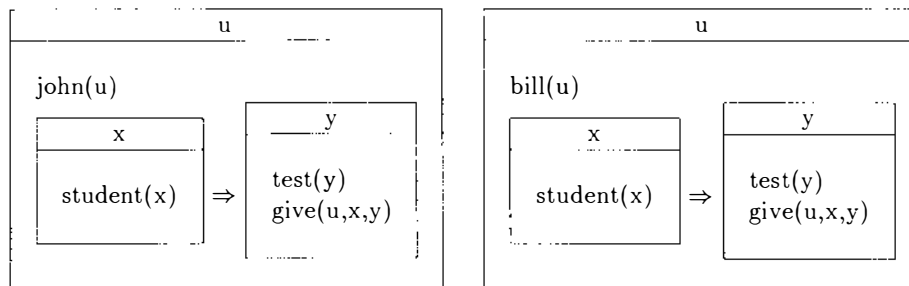
3.1 Parallel Scope Examples

Consider example (1), repeated here:

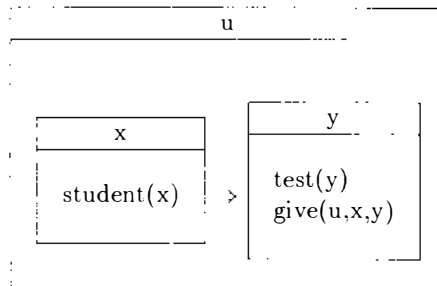
- (5) John gave every student a test, and Bill did too.

We consider a parallel reading, in which *a student* takes wide scope in both conjuncts, and a non-parallel reading, in which *a student* takes wide scope in the first conjunct and narrow scope in the second.

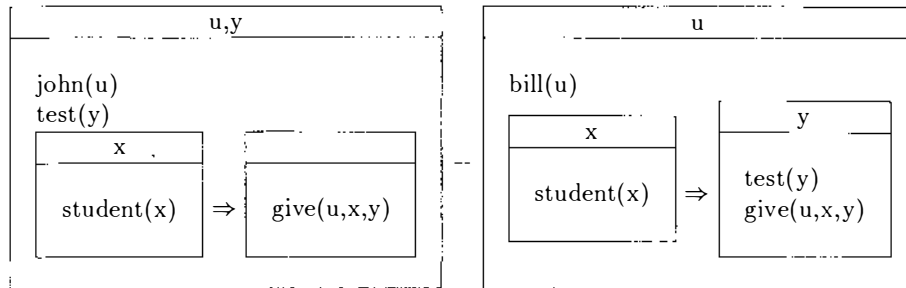
Parallel reading:



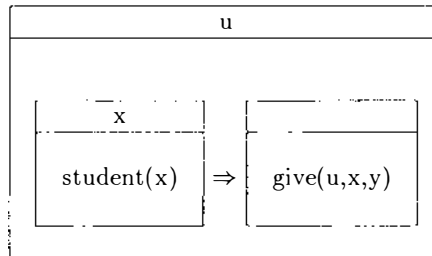
This produces the following maximal common theme, T_1 :



Non-parallel reading:



This produces the following maximal common theme, T_2 :

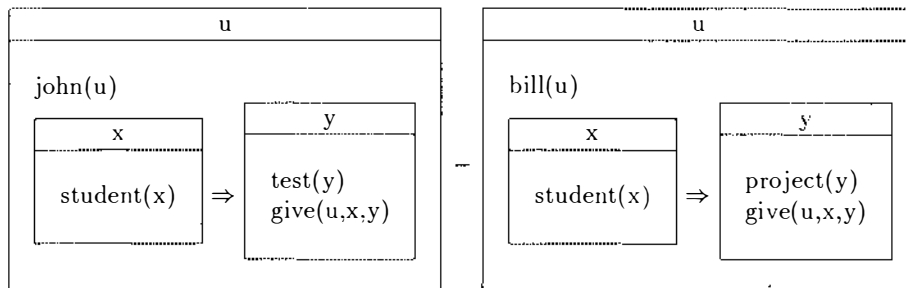


Since $T_1 \rightsquigarrow T_2$, the parallel reading is preferred over the non-parallel reading. Similar reasoning will apply to the other pair of parallel and non-parallel readings.

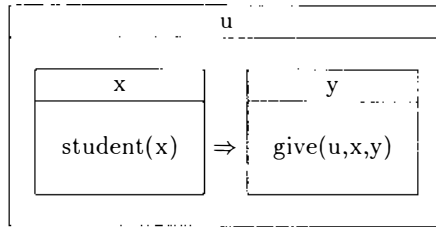
Consider now (2), repeated here:

- (6) John gave every student a test, and Bill gave every student a project.

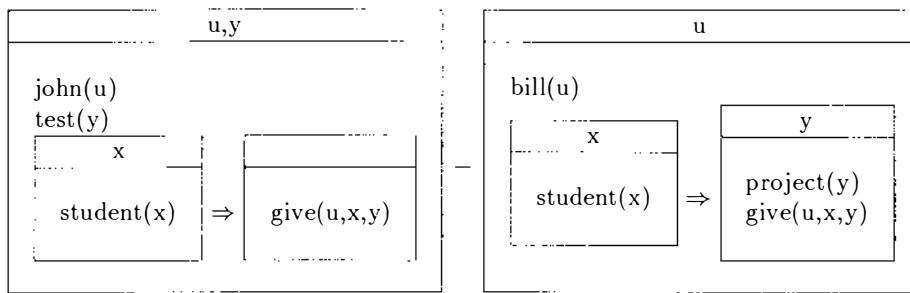
Parallel reading:



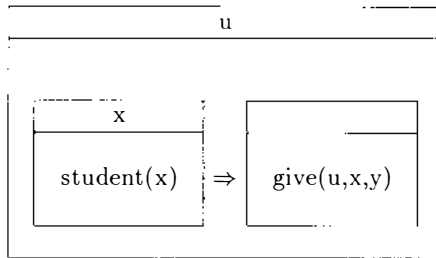
This produces the following maximal common theme, T_1 :



Non-parallel reading:



This produces the following maximal common theme, T_2 :



Again, we have $T_1 \rightsquigarrow T_2$.

Our account would apply in a similar way to the following examples, which exhibit the same preference for parallel scoping:

- (7) a. John gave every student a project, and Bill gave every student an assignment, too.
 b. John gave every linguistics student a project, and Bill gave every philosophy student a test.

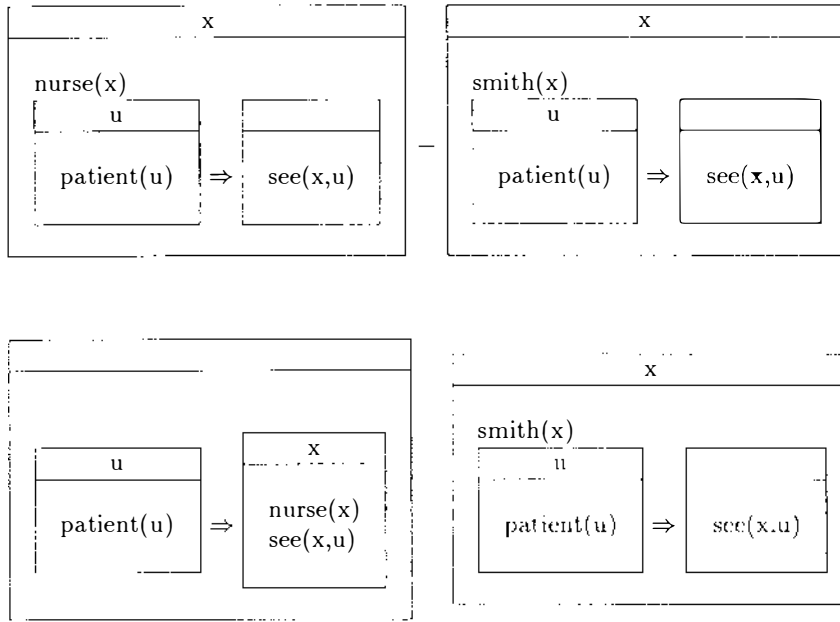
4 Wide-scope Puzzle

We now examine the wide-scope puzzle first observed by (Sag, 1976), illustrated by (3), which is repeated here:

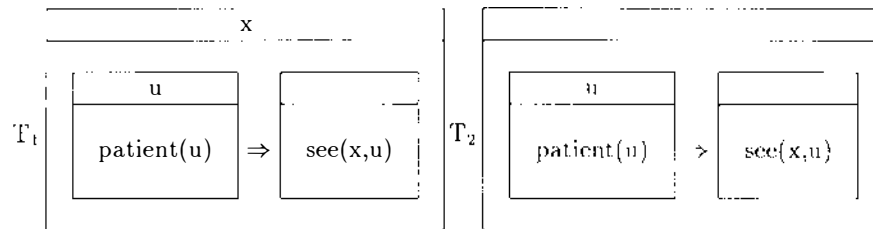
- (3) A nurse saw every patient. Dr. Smith did too.

Sag observed that, while the first sentence in isolation would have the expected two possible readings, in which *a nurse* can take either wide or narrow scope, (3) only permits one reading, in which *a nurse* must take wide scope. We now show that this is a consequence of our *Maximization Constraint*.

There are two potential readings for (3):



The Maximal Common Themes for these two readings are:



Since $T_1 \rightsquigarrow T_2$, the wide scope for *a nurse* is preferred.

5 Related Work

5.1 Related Approaches

There are many ellipsis accounts (eg, (Sag, 1976; Williams, 1977; Dalrymple et al., 1991; Fiengo and May, 1994)) that capture certain scope parallelism effects that occur in ellipsis contexts. However, as we have argued above, these accounts do not apply to similar effects where ellipsis does not occur. Also, none of these approaches successfully account for the wide scope puzzle. There are also discourse accounts, such as that of (Prüst et al., 1987), which are similar in spirit to ours. But it is not clear how an approach like Prüst’s would apply to the data examined in this paper. In particular, it does not apply to the wide scope puzzle.

However, the approach of (Fox, 1995), does provide an account of the wide-scope puzzle and other data we have considered above, and we now turn to an examination of this approach. In this account, a general parallelism constraint (Rooth, 1992; Fiengo and May, 1994)⁸) together with an economy constraint on Scope Shifting Operations captures scope effects, including the wide-scope puzzle.

Consider example (3) again:

- (3) A nurse saw every patient. Dr. Smith did too.

Fox presents the following *ellipsis scope generalization* (ESG):

the relative scope of two quantifiers...may differ from the surface c-command relation only if the parallel difference will have semantic effects in the elided VP.

By the ESG, *every patient* cannot QR over *a nurse* because it would not have a semantic effect for *every patient* to QR over *Dr Smith*. In Fox’s theory, ESG is a consequence of parallelism plus “economy”, which prohibits an application of QR (and related operations) if it does not have a semantic effect. There are two fundamental differences between Fox’s approach and ours, the first theoretical, and the second empirical.

We consider first the theoretical difference. In our approach, the wide-scope puzzle is captured by the parallelism theory – we don’t need to appeal to economy as well. The fact that we are applying parallelism to DRS’s ensures that no quantifier can outscope a name, because in DRT discourse referents and conditions introduced by names must occur in the top level DRS (see (Kamp and Reyle, 1993) for details).⁹ For Fox, the inability of a quantifier to outscope a name in subject position is a consequence of economy, which doesn’t permit an operation to apply unless it has a semantic effect. The DRT processing constraint on names might be construed as a principle of economy in some sense, since it eliminates certain spurious ambiguities. But the economy principle employed by Fox is quite different from this. We draw attention to two crucial differences. First, Fox’s constraint incorporates a version of the “Have

an Effect on Output Condition” constraint into the derivation system, which introduces major conceptual and computational complications (see (Johnson and Lappin, 1997) for discussion). Second, Fox’s approach complicates the syntax/semantics interface, since it requires information about possible readings in determining the applicability of a syntactic operation. The DRT rule for representing names does not involve any of these complications.

We now turn to empirical differences. In our approach, the wide-scope puzzle reflects a parallelism in representation between *names* and *wide scope existential quantifiers*. For Fox, the effect is much more general and leads to empirical differences. In the following examples, Fox predicts a required wide scope subject reading, while our approach permits both readings:

- (8) A nurse saw every patient. Every doctor did too.
- (9) Every student read a book and Harry did too.

We would argue that the wide-scope effect is not present here; the narrow scope object reading is permitted. This is particularly clear in (9), where we find that the reading in which *a book* takes wide scope over *every student* is perfectly acceptable.

Conversely, our approach enforces the wide-scope subject preference in the following example, while Fox does not:

- (10) Dr. Smith saw every patient. A nurse did too.

Here, *every patient* can take wide scope over *a nurse*. (p. 334, ex 79). Our approach applies symmetrically, treating this example just like (3).

5.2 Problematic Examples

Several researchers have called into question the reality of the “wide-scope puzzle”. (Johnson and Lappin, 1997) present the following apparently problematic examples:

- (11) At least one Labour MP attended most committee meetings, and Bill did too.
- (12) At least one natural number other than one divides into every prime number, and one does too.
- (13) At least two cabinet members bear responsibility for each government department, and the Prime Minister does too.

(Johnson and Lappin, 1997) observe that the subject NP’s in these sentences can take narrow scope, despite the occurrence of a name in parallel position in the second sentence¹⁰. While these are delicate judgements, we concur that the wide-scope effect does not appear to be present in these cases.

Our account imposes a preference ordering on possible DRS representations. If the preferred representation conflicts with general knowledge or expectations, it may be rejected in favor of some other representation. This provides a reasonable explanation for (12), since the wide scope reading conflicts with general knowledge about numbers. However, it is not clear to us that the wide scope readings for examples (11) and (13), conflict with general knowledge (at least for American English speakers). Consequently, we will explore the possibility that a structural effect accounts for the absence of the wide scope reading in these cases.

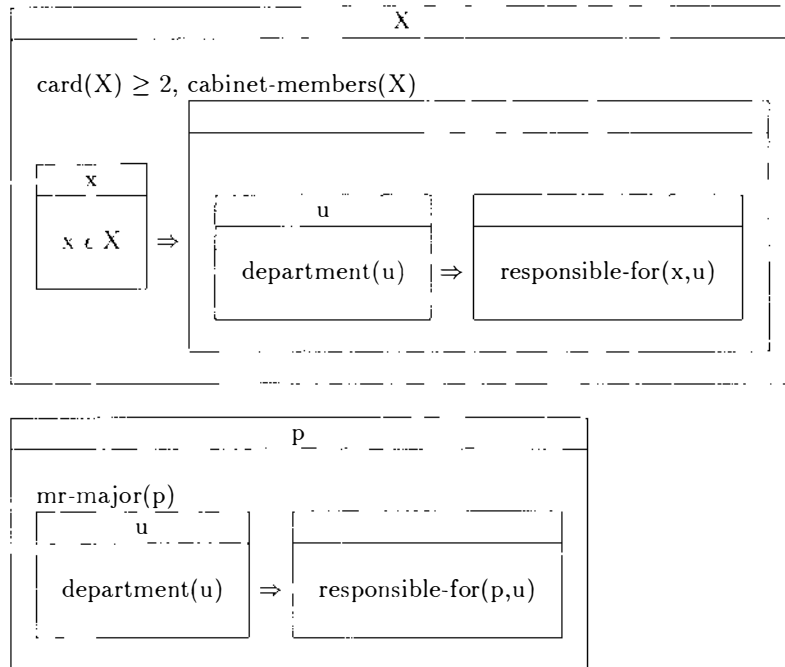
We will examine the following variant of example (13):¹¹

- (14) At least two cabinet members bear responsibility for each government department, and Prime Minister Major does too.

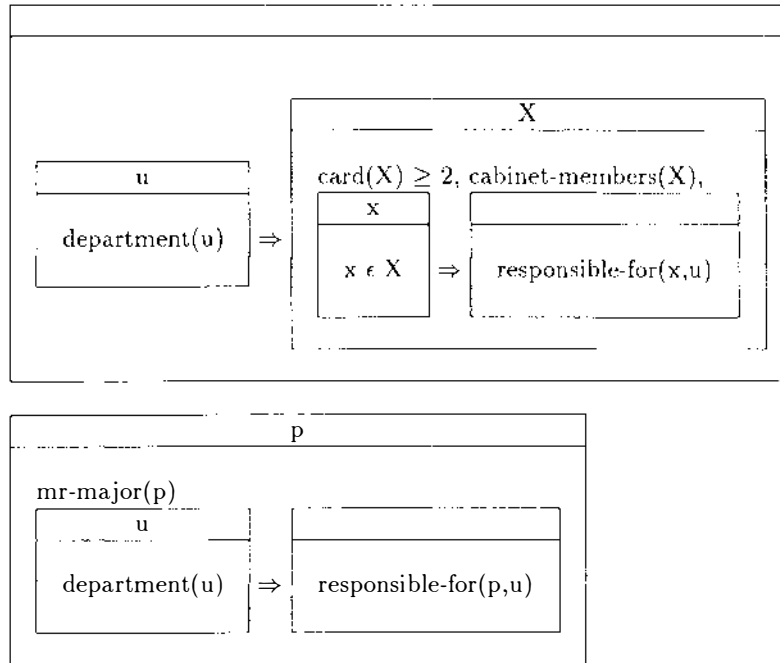
The NP *at least two cabinet members* introduces a plural discourse referent. The most natural reading is a distributive one – each of the cabinet members mentioned bears responsibility for each department, rather than the responsibility being a collective one. Thus, following (Kamp and Reyle, 1993, page 327), we introduce a duplex condition, or quantificational structure, representing the distribution.

Thus we have the following representations for the two readings of (14):

Reading 1:



Reading 2:



On either reading, the two sentences do not share the same nesting structure. Thus, no Maximal Common Theme can be constructed, and our approach does not impose a scope preference. For (12) and (11), we would give similar representations, based on the idea that any NP of the form *at least N* can be represented as a plural discourse referent with a quantificational structure representing a distributive reading.¹²

While the judgements are delicate, we find that the wide-scope effect reappears in example (13) if the determiner “at least two” is replaced by a simple indefinite “a”, and we find a similar effect with (11). However, a replacement in example (12), does not reinstate the wide-scope effects. This supports our view that (11) and (13) exhibit a structural effect, while (12) must be explained in terms of knowledge about numbers.

6 Extending the Approach

In this section, we discuss some extensions to the approach, both to “loosen” the constraints on parallelism in certain ways, and to capture similar effects in embedded contexts.

6.1 *Loosening the Parallelism Constraint*

In the account described so far, scope parallelism effects are captured in examples where the related DRS's have the same nesting structure, and identical parallel quantifiers. In this section, we suggest some ways in which we envision "loosening" our account to capture some parallelism effects with non-identical quantifiers and differences in nesting structure.

The following example involves different quantifiers in parallel positions:

- (15) John gave every student a test. Bill gave most students a project.

In our view, the relative scope between *a test* and *every student* will be the same as the relative scope between *a project* and *most students*. To capture this, we envision adding an operation like the following:

Replace a quantifier Q with a quantifier Q' where either both Q and Q' are monotone increasing in their second arguments or both Q and Q' are non-monotone increasing in their second arguments

This makes it possible to replace *every* with *most* in the above example. The following example illustrates a difference in nesting structure:

- (16) A nurse saw every patient, but Dr. Smith didn't.

The negation in the second clause introduces an additional level of nesting, making it impossible to enforce parallel scoping effects. To capture this, we propose to ignore negation for the purpose of determining MCT. More generally, we will explore an extension of our approach in which we allow determination of MCT to proceed with respect to Modified Embedding Trees (Asher, 1993), in which some levels of embedding can be "collapsed".

6.2 *Embedded Readings*

In all the examples we have considered, the two related discourse constituents are at the top level of the DRS. The situation becomes more complicated if the related constituents are embedded. Consider the following example:

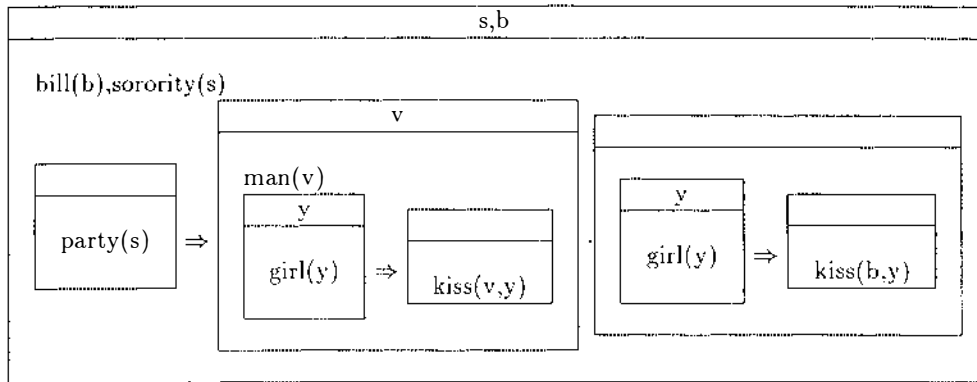
- (17) If the sorority has a party, a man will kiss every girl at the party and Bill will too.

We believe the same effect occurs here – because of the parallelism of *Bill* and *a man*, there is a strong preference for *a man* to take wide scope with respect to *every girl*.

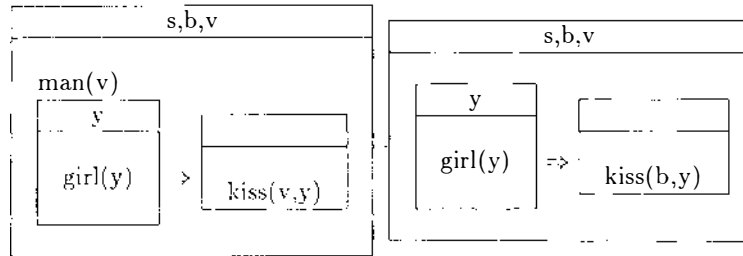
As currently formulated, our approach would not capture the scope parallelism effects in embedded contexts. However, a simple modification of the approach will accommodate this. The related DRS's will contain a list of all *accessible* discourse markers. For a DRS occurring on the top-level, that is just

those discourse markers introduced in that DRS. However, in an embedded DRS, this will expand the list of discourse markers.¹³ Consider our representation of (17):

Parallel Reading:

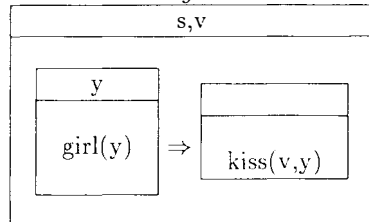


Here, we have two related DRS boxes constituting the consequent of the conditional, representing, respectively, *A man will kiss every girl* and *Bill will kiss every girl*. Although the discourse marker *b* for *Bill* is introduced at the top level, in terms of our parallelism determination it also appears on the discourse marker list of the embedded DRS's. Thus we represent the related constituents as follows:



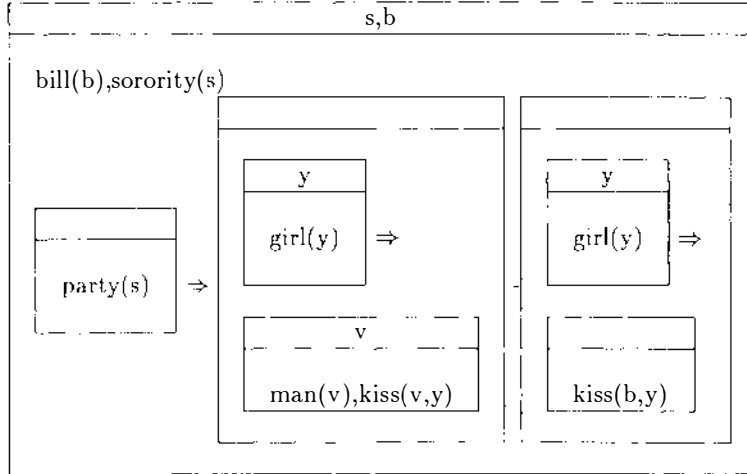
The following represents the MCT constructed for these two constituents¹⁴:

Parallel Reading MCT:

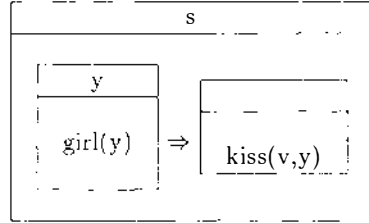


Consider now the non-parallel reading, where *a man* takes narrow scope with respect to *every girl*:

Non-Parallel Reading:



Non-parallel Reading MCT:



In this case, the discourse marker v (representing the “kisser”) does not appear at the top level, as it does in the MCT for the parallel reading. Since the Parallel MCT \rightsquigarrow the Non-Parallel MCT, the preference is accounted for.

7 Conclusions

We have defined a general mechanism for imposing a preference for parallel structure in related discourse constituents, using the SDRT framework. We have shown that this captures parallel scope effects both in ellipsis and non-ellipsis contexts. In addition, because of the DRT treatment of names, our parallelism mechanism also captures puzzling effects involving wide scope readings.

Endnotes

^{0*} The second author would like to acknowledge support by a National Science Foundation CAREER grant, IRI-9502257.

¹This observation figures prominently in most of the ellipsis literature, going back at least to (Sag, 1976), and (Williams, 1977).

²(Tancredi, 1992) also observes that similar interpretive effects arise whether or not ellipsis is present, and attempts to define a mechanism general enough to account for this. However, Tancredi does not examine issues of scope parallelism, but rather looks at the strict/sloppy alternation.

³(Kitagawa, 1991) describes a VP copy operation which can precede indexation, although he does not discuss scope determination. Also, (Fox, 1995) observes that QR and the VP-copy operation could be freely ordered at LF.

⁴One example of such an account is that in (Lappin, 1984). In that account, a Cooper storage mechanism is used to obtain wide-scope readings. Thus, the recovered material for VP ellipsis is a pair consisting of VP matrix and quantifier store. A quantifier can thus be discharged from store either before or after being recovered at the ellipsis site.

⁵We represent DRS's using the lambda operator together with the ordinary "box" notation of DRT. This makes it possible to describe a straightforward compositional DRS construction algorithm. See (Muskens, 1996) or (Asher, 1993) for discussion.

⁶(Prüst et al, 1987) define a mechanism for computing the *most specific common denominator* (MSCD) of related discourse utterances. This is in much the same spirit as our Maximal Common Theme. However, Prüst et al do not attempt to account for the sort of phenomena discussed in this paper.

⁷In fact, this operation is not needed to capture the facts considered in this paper. We leave as a topic of future research the question of whether a generalization operation is relevant to determination of parallelism effects.

⁸(Rooth, 1992) and (Tancredi, 1992) define constraints to the effect that two related clauses are identical modulo the focused elements. This is used as a constraint on possible readings in VPE.

⁹(Muskens, 1996) has objected that the DRT construction rule for names makes a compositional approach difficult. While this is true for the approach of (Kamp and Reyle, 1993), we feel there are a variety of ways to address this issue, such as an approach using underspecified conditions like that of (Reyle, 1993). Indeed, our argument concerning the wide-scope puzzle could be viewed as providing additional support for the DRT construction rule for names.

¹⁰(Fox, 1995) argues that (11) permits ambiguity because of an existential quantification over an event variable. However, this explanation will not apply to (12) and (13), where there is no event variable. See (Johnson and Lappin, 1997) for further discussion.

¹¹We replace the definite description *the Prime Minister* with a name, *Prime Minister Major*, since, in our approach, a definite description may appear either with wide or narrow scope, although in this example the narrow scope reading would be equivalent to the wide scope reading.

¹²This is consistent with the view in (Szabolsci, 1997) that quantifiers involving a modified numeral are distinct from ordinary indefinites. We also represent *at least one* as a duplex structure in our analysis of (11). *at least one* appears to have two readings, a quantificational one (Hans Kamp, p.c.), and a specific indefinite reading that appears to arise only when a singular pronoun has an NP containing *at least one* as an antecedent. The latter reading is not relevant to our examples.

¹³The reader is invited to verify that this modification of DRSs does not alter any of the earlier predictions we have made.

¹⁴As the reader can verify, this was constructed by applying the following operations: (to the first constituent) *delete man(v)*, *delete discourse marker b*; (to the second constituent) *delete discourse marker v*, *rename b as v*.

References

- Nicholas Asher. 1993. *Reference to Abstract Objects in English*. Dordrecht.
- Mary Dalrymple, Stuart Shieber, and Fernando Pereira. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy*, 14(4), August.
- Robert Fiengo and Robert May. 1994. *Indices and Identity*. MIT Press, Cambridge, MA.
- Danny Fox. 1995. Economy and scope. *Natural Language Semantics*, 3(3):283–341.
- Hub Prüst, Martin van den Berg, Remko Scha. 1987. A discourse perspective on verb phrase anaphora. *Linguistics and Philosophy*, 17(3):261–327.
- David Johnson and Shalom Lappin. 1997. A critique of the minimalist program. *Linguistics and Philosophy*.
- Hans Kamp and Uwe Reyle. 1993. *From Discourse to Logic*. Kluwer, Dordrecht.

- Hans Kamp. 1980. A theory of truth and semantic representation. In J. Groenendijk, T.M.V. Janssen, and M. Stokhof, editors, *Formal Methods in the Study of Language*, pages 277–322. Dordrecht. Volume 136.
- Yoshihisa Kitagawa. 1991. Copying identity. *Natural Language and Linguistic Theory*, 9(3):497–536.
- Shalom Lappin. 1984. VP anaphora, quantifier scope, and logical form. *Linguistic Analysis*, 13(4):273–315.
- Alex Lascarides and Nicholas Asher. 1993. Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy*, 16:437–493.
- Robert May. 1977. *The Grammar of Quantification*. Ph.D. thesis, MIT.
- Reinhard Muskens. 1996. Combining Montague Semantics and Discourse Representation. *Linguistics and Philosophy*.
- Uwe Reyle. 1993. Dealing with ambiguities by underspecification: Construction, representation and deduction. *Journal of Semantics*, 10(2).
- Mats Rooth. 1992. Ellipsis redundancy and reduction redundancy. In *Proceedings of the Stuttgart Ellipsis Workshop*, Stuttgart, Germany.
- Ivan A. Sag. 1976. *Deletion and Logical Form*. Ph.D. thesis, Massachusetts Institute of Technology. (Published 1980 by Garland Publishing, New York).
- Anna Szabolsci, ed. 1997. *Ways of Scope-Taking*. Kluwer Academic Publishers, Dordrecht.
- Christopher Tancredi. 1992. *Deletion, Deaccenting and Presupposition*. Ph.D. thesis, Massachusetts Institute of Technology.
- Edwin Williams. 1977. Discourse and logical form. *Linguistic Inquiry*, 8(1):101–139.