The Logical Foundations of Discourse Interpretation

Nicholas Asher

A central but still unsettled question in formal theories about discourse interpretation is: What are the key theoretical structures on which discourse interpretation should depend? If we take our cue from theories that analyze the meanings of individual sentences, the meaning of the discourse's parts should determine the meaning of the whole; some sort of principle of compositionality of meaning must hold at the level of discourse interpretation. So a theory of discourse interpretation must develop from an account of discourse structure.

Unlike the syntactic structure of a sentence, the discourse structure of a text is not a structure studied by syntacticians or governed principally by syntactic concerns. It has to be inferred from a variety of knowledge sources. Recent work on discourse structure in AI, philosophy and linguistics has shown that discourse structure depends on numerous information sources—compositional semantic principles, lexical semantics, pragmatic principles, and information about the speaker's and interpreter's mental states. So a theory of discourse interpretation must in fact also be a theory of semantics and pragmatics and their interaction—a theory of the pragmatics-semantics interface.

Such a theory linking together pragmatics and semantics brings up a foundational question about frameworks. Pragmatics, though not often formalized, has often made appeal to different types of logical principles than semantics. While semantic theories have typically used a classical, monotonic, logical framework, pragmatic theories appear to best couched within nonmonotonic logic. How should we model the interaction of these multiple knowledge sources needed to construct discourse structure, or the interaction between defeasible pragmatic principles and nondefeasible semantic principles?

---

[1]Received June 20, 1996; revised version Marsh 19, 1997.

Another fundamental difference between pragmatics and semantics also requires resolution. The model-theoretic approach to semantics is thoroughly entrenched, whereas in pragmatics appeals are often made to representations of information, beliefs and other mental states of the participants. How should a discourse context be thought of—as a structured representation, or model-theoretically? How should we model in a logically precise fashion the updating of discourse contexts with new information?

One of the good things about recent developments in theories of discourse interpretation is that we know at least some of the answers to these questions. I will try to provide a guide to some strategies for building such theories. I will begin with a simple and now standard semantic approach to discourse interpretation, Discourse Representation Theory (DRT) without eventualities. I will then add in steps components needed for pragmatic interpretation. The approaches become more complex, but each bit of additional complexity lets us analyze some phonenomenon or represent some bit of pragmatic reasoning that we were not able to before. The end result will be a theory of the semantic-pragmatic interface, Segmented Discourse Representation Theory or SDRT. Like DRT, SDRT exploits representations to model discourse contexts and to determine the semantic effects of discourse structure.

Finally, a question that should be asked in this context, is why should any of this be of interest to logicians? Like formal semantics, discourse semantics is an area of application for logical techniques. The few theorems mentioned in this text are not terribly surprising nor do they introduce any new formal techniques unfamiliar to logicians. But discourse semantics has uncovered some new logical problems that were not part of standard formal semantics and could not easily be formulated within the framework of higher order intensional logic, the logical framework of Montague Grammar. It is my hope that some of these problems and the attempts that semanticists have made to formalize them will be of interest to logicians and may prompt them to investigations in these areas.

1 Motivations for Dynamic Semantics

To understand the development of dynamic semantics, one must understand the problems that motivated this development and to do that, one must say something about the interpretation of the constituents of discourse —its constituent sentences and their constituent phrases— i.e., standard formal semantics. In the 1960s Montague developed a very influential theory for analyzing the truth conditions of sentences, Montague Grammar. The question then arose: How do we now analyze the truth conditions of a discourse? For just as one can only understand the meanings of expresions within the context of a sentence, the meaning of a sentence is not something that occurs in isolation—there is always a discourse context.

One might think that there is nothing more to the interpretation of discourse than simply building up the meanings of its constituent sentences and then combining them together. If meanings are taken to be sets of possible worlds or other indices, the operation of combination is particularly simple; it's intersection. But once a clear and precise proposal for the semantics of sentences was given by Montague Grammar, several problems emerged that showed that the interpretation of discourse would not be simple. Below I discuss two: pronominal anaphora and the interpretation of tense in discourse.

1.1 Pronominal Anaphora

Pronominal anaphora occurs when an anaphoric pronoun refers back to some word or phrase in the preceding discourse. In the second sentence of the discourse (1) below, the anaphoric pronouns *he* and *it* refer back respectively to the farmer and the donkey.

(1) A farmer owns a donkey. He beats it.

In translating each sentence of (1) into logical formulae using the tools of Montague Grammar, we encounter a problem in the translation of anaphoric pronouns. The pronouns in the second sentence are anaphorically bound to the noun phrases in the first sentence, and perhaps the most natural way to try to represent this linkage in the logical form of (1) is to translate the pronouns as variables that are to be bound by the quantifiers introduced by the noun phrases in the first sentence. But the problem is that by the time we attempt this "bound variable" translation of the pronouns, we have finished the translation of the first sentence and so closed off the rightward scope of the quantifiers. Conjoining the translation of the second sentence does not produce a bound variable reading of the variables introduced by the pronouns. In fact it produces the open sentence:

(1') $\exists x(farmer(x) \wedge \exists y(donkey(y) \wedge owns(x, y))) \wedge beats(x, y)$

The Montague Grammarian's approach to quantifying-in offers a partial solution to this problem (Gamut 1991). However, if we wish to use the procedure of quantifying in to deal with discourses in which anaphoric linkage to an antecedently occurring noun phrase exists over multiple sentences as in (2) below, then we must suppose a complete syntactic analysis of the discourse prior to the interpretation of any of its constituent sentences.

(2) A farmer owned a donkey. He beat it. It ran away.

This conclusion is not cognitively plausible and prompts us to look for a different solution to the problem of intersentential anaphora.

1.2 Temporal Anaphora

Observations of Barbara Partee (1973) and of Kamp and Rohrer working on the analysis of French tenses in discourse during the seventies brought to attention a facet of the meaning of tenses that was missing from the best analyses of tense of the day. Those analyses took tenses to be tense operators of the sort found in tense logic (cf. Montague 1974, and some of the references in Dowty 1979). In a French discourse like

(3) Pierre entra dans le salon. Il s'assit sur la banquette. Il s'endormit.
Pierre entered the room. He sat down on the sofa. He fell asleep.

the three events introduced, Pierre's entering, his sitting down and his falling asleep, all occur in the past if (3) is true, but they also occur in a definite sequence—the sequence in which they are introduced in the discourse. That is, we naturally understand the story as telling us that his entering the room occurred prior to his sitting down which in turn occurred prior to his falling asleep. The operator view of tenses, on which a past-tensed sentence the logical form of which is $P\phi$ is true iff ϕ holds of some time prior to the moment of speech, is incapable of capturing this contextual sensitivity. Even views on which verbs are treated as predicates of events and the past tense introduces a relation of earlier than between the event and the moment of speech are not by themselves equipped to capture this context sensitive interpretation.

2 Dynamic Semantics and Basic DRT

The solution that Kamp and Heim independently proposed to the problem of anaphoric pronoun interpretation was to redefine the semantic contribution of a sentence and its constituents to a discourse. This was the first attempt at discourse semantics, using rigorous, formal methods similar to those found in Montague Grammar. In Montague Grammar and on other accounts of discourse interpretation (Stalnaker 1978), the contribution of a sentence is a proposition, or, formally, a set of possible worlds in which the sentence was true. Such a proposition contributes to the content of a discourse in a simple way: the meaning of a discourse is just those possible worlds that are in the intersection of all the propositions that are the meanings of the discourse's constituent sentences. For Kamp and Heim, a sentence S, when interpreted in a discourse D no longer simply yields a set of worlds; the meaning of S is rather a relation, a relation between contexts. This new relational conception of the meaning of S was dubbed its context change potential.[2]

To define the context change potential of a sentence, Kamp uses a representational theory of discourse semantics, Discourse Representation Theory (DRT). DRT assigns a truth conditional meaning to a natural language discourse in two steps: the DRS construction procedure and the correctness definition. In the first step, we construct a representation of the content of the discourse known as a discourse representation structure or DRS. DRT uses DRSs to define context change potential. I won't detail here how DRSs for clauses and their interpretation can be built up compositionally (for details see Asher 1993 or Muskens 1996); I want to focus on discourse aspects of dynamic semantics.

The basic fragment of DRT is defined by the following definition of DRSs and DRS conditions.

Discourse Referents is a set of objects denoted by x, y, z, with or without subscripts. Predicates is a set of predicate constants associated with various natural language nouns, verbs and adjectives. Supose $U \subseteq$ Discourse Referents; we then define DRSs K and conditions γ recursively:

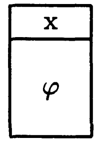
$$K \quad := \quad \langle U, 0 \rangle \mid K^\frown \gamma$$

Let $R \in$ Predicates be an n-ary predicate and $x_1, \cdots, x_n$ be discourse referents.

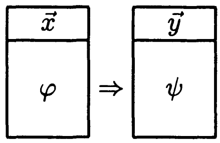
$$\gamma \quad := \quad R(x_1, \cdots, x_n) \mid \neg K \mid K_1 \Rightarrow K_2 \mid K_1 \vee K_2.$$

The truth definition for discourses that give rise to DRSs described by the definition above is given by embedding the DRSs they generate into a standard Tarskian model. Given the semantics for DRSs, any DRS in the fragment above has a first order translation. E.g.

[2]Many other formalisms have adopted the Kamp-Heim approach in a different guise. See Barwise (1985), Gronendijk and Stokhof (1987).



has a proper embedding in a model $M = \langle D, A_1, \cdots, A_n, \cdots \rangle$ (where the A_i's represent the extensions of the non-logical predicate symbols of the DRS language) relative to some embedding g iff $\exists x \phi$ is is satisfied in M relative to some assignment to free variables. Similarly,



is satisfied relative to some embedding function g in M iff $\forall x \exists y(\phi \rightarrow \psi)$ is satisfied relative to some assignment.

Let us now make this more precise with a full semantic definition of proper embedding. I define simultaneously the model theoretic transition P and the satisfaction of conditions V relative to a model.

Definition 1:

$$fP_M(U,O)g \quad \textit{iff} \quad f \subseteq g \wedge dom(g) = dom(f) \cup U$$

$$f \in V_M(R(x_1, \cdots, x_n)) \quad \textit{iff} \quad R_M(f(x_1), \cdots, f(x_n))$$

$$fP_M(K^\frown\gamma)g \quad \textit{iff} \quad fP_M(K)g \wedge g \in V_M(\gamma)$$

$$f \in V_M(\neg K) \quad \textit{iff} \quad \neg\exists g\, fP_M(K)g$$

$$f \in V_M(K \Rightarrow K') \quad \textit{iff} \quad \forall g\,(fP_M(K)g \rightarrow \exists\, h\, gP_M(K')h)$$

$$f \in V_M(K \vee K') \quad \textit{iff} \quad \exists\, g\, fP_M(K)g \vee \exists h\, fP_M(K')h$$

I assume a sufficiently large set MOD of models, viz. those formed from maximal consistent saturated sets of first order formulas.

In order to formulate an insightful correspondence between DRT and first order logic, it is useful not only to define P but also its *lifted* counterpart $\mathcal{P}$ on sets of model sequence pairs (MSP)'s for a DRS K. $\mathcal{P}(K)$:Pow(MSP) $\rightarrow$ Pow(MSP) is defined distributively or pointwise over the set of model sequence pairs , exploiting K's effect on an input MSP (M, g) to produce a certain output MSP (M', g'), where . That is for a context X:

$$[\mathcal{P}(K)](X) = \{(M', g') : \exists(M, g) \in X(M = M' \wedge gP_M(K)g')\}$$

One can associate with each DRS K both a set of pairs $\langle M, f \rangle$ where f is a proper embedding of K in a Tarskian model M. Such sets are called information states and they correspond to a a first order formula. This connection is made precise in the following lemma (proved by induction). I use here the formulation of Fernando (1994). Let $\{\langle w, 0 \rangle : M \in MOD\} = \sigma_0$
Characterization Lemma for DRT:

- For every first order formula χ with a set of free variables U, there is a DRS (U, C) such that $\mathcal{P}(U, C)[\sigma_0] = \{(M, f) : Dom(f) = U$ and $M \models \chi[f]\}$
- Every DRS (U, C) has a characteristic formula χ where U is the set of free variables in χ and $\mathcal{P}(U, C)[\sigma_0] = \{(M, f) : Dom(f) = U \text{ and } M \models \chi[f]\}$

DRT is a dynamic theory of discourse interpretation. The idea is that each sentence has a *context change potential* (CCP) that changes a discourse context when the information contained within the sentence is added to the context. The background context is represented as a DRS; the contribution of the individual sentence is the context DRS extended with the conditions and discourse referents contributed by the processing of the sentence. Thus, the CCP of a sentence is a relation between DRSs. More precisely, let C_ϕ be the set of conditions of the DRS ϕ, χ_ϕ be the characteristic formula of ϕ, and let

$$\phi^\frown\psi = ((U_\psi \cup U_\phi), (C_\psi \cup C_\phi)), \textit{ if } (\chi_\phi \wedge \chi_\psi) \textit{is first order consistent.}$$

The CCP of an unambiguous sentence S, which by a procedure known as the DRS construction procedure yields a DRS K_S,[3] can now be represented as the DRS transition predicate T_{drs}, between consistent DRSs. We define T_{drs} as the set

$$\{(K, K_S, K^\star) : K \textit{is a DRS, } K_S, K^\star \textit{are consistent DRSs and } K^\star = K^\frown K'\}$$

[3]For details see Kamp and Reyle (1993).

The CCP of a sentence may also be represented as a relation between information states (assuming once again the presence of the DRS construction procedure). A model-theoretic context is a set of model sequence pairs (MSPs), and so a natural candidate for the model-theoretic notion of the CCP of a sentence S is just the lifted function P applied to the DRS derived from S.

Using the Characterization Lemma, Fernando (1995) proves a precise equivalence between the representational and model-theoretic conceptions of CCP for the simple core fragment of DRT and he has also shown that the notion of equivalent CCP is r.e. The representational level and the model theoretic level yield bisimulation-equivalent notions of context and CCP.[4] Let $\Phi_\perp$ be the set of absurd formulas, and let MOD be defined as above). Define Acc as the smallest set of states such that, where $[\mathcal{P}(\phi)](\sigma)$ represents the application of the function $\mathcal{P}(\phi)$ to σ:

$$\sigma_0 \in Acc \wedge (\sigma \in Acc \rightarrow [\mathcal{P}(\phi)](\sigma) \in Acc)$$

Thus, we can define within Acc for any consistent $\phi, \sigma_\phi = [\mathcal{P}(\phi)](\sigma_0)$. Suppose that $\mathcal{P}(\phi^\frown\psi) = \mathcal{P}(\phi) \circ \mathcal{P}(\psi)$ and define $\phi \underline{\leftrightarrow} \psi$ iff $\forall \phi' (\phi^\frown\phi' \in \Phi_\perp \leftrightarrow \psi^\frown\phi' \in \Phi_\perp)$.

Theorem 2.1 *(Fernando 1995):* $\phi \underline{\leftrightarrow} \psi$ iff $\sigma_\phi = \sigma_\psi$*; and further,* $\underline{\leftrightarrow}$ *and* $\Phi_\perp$ *are r.e.*

$\underline{\leftrightarrow}$ defines a bisimulation relative to the function $^\frown$, and $\mathcal{P}(K)[\sigma_0]$, in effect, exploits $\underline{\leftrightarrow}$ to induce a bisimulation relative to state transitions, the model-theoretic interpretation of a DRS on a set of MSP's. Further, bisimilarity on state transitions is strongly extensional; bisimilar transitions on states have identical outputs when applied to the empty information state, which explains the first part of the theorem. For the second part of the theorem, note that $\underline{\leftrightarrow}$ is defined proof theoretically, by an operation on DRSs. Given the *Characterization Lemma*, every DRS is equivalent to a first order formula. Hence, the notions $\underline{\leftrightarrow}$ and $\Phi_\perp$ are first order definable and so r.e.

The effect of this theorem is to show that the model-theoretic and the representational conceptions of CCP in dynamic semantics are equivalent. Far from "eliminating representationalism," the inclusion in MSPs of functions from discourse referents to objects makes them essentially representational;

---

[4]A similar representation theorem for a slightly different conception of model theoretic context is sketched in Asher (1993). For more on the notion of bisimulation, see Aczel (1989) and Park (1988).

this information is not about the world but about how our information about the world is structured.[5] But using MSPs is useful because they indicate the minimal amount of representational information needed to treat phenomena for which DRT was designed. I apply this strategy now to pragmatic-semantic contexts.

3 DRT_1—DRT with events

Without events or temporal constants of any kind, the basic DRT fragment is unable to make sense of change. In the timeless models of DRT_0, the following simple discourse is (unintuitively) inconsistent:

(4) John sat down. John got up.

To analyze change, DRT follows a Davidsonian approach and introduces eventualities as additional arguments of conditions derived from verbs. The only change to the definition of DRSs and DRS conditions from section 1 is to add an eventuality type discourse referents and to change the set of Predicates— those derived from verbs will have one more argument place than before. This makes possible a satisfactory analysis of change. I will here introduce ways of treating tense that make the discourse (4) consistent, though they do not give a complete temporal interpretation of tensed discourse. Much of the temporal interpretation of discourse is pragmatically determined. The notion of CCP equivalence remains r.e. for DRT_1.

DRT_0 and DRT_1 define distinct consequence relations for some natural language discourses by offering a translation into a formal language that is given a model theoretic interpretation. So for instance, a discourse like (4), when translated (via the DRS construction procedure) as a DRS ϕ in DRT_0, is inconsistent, provided we make certain assumptions about the relationship between the predicate sit down and get up (viz. one cannot simultaneously get up and sit down). However, the translation of (4) is consistent in DRT_1. We could represent this as (4) $\vdash_0 \bot$ but $\not\vdash_1 \bot$.

[5]Many authors who have rejected the representationalism of DRT do not seem to acknowledge that the use of assignment functions in their semantic values amounts to an equal representational commitment.

4 Beyond Dynamic Semantics

DRT, and dynamic semantics in general, make an important contribution to our understanding of how the discourse context affects pronominal and temporal interpretation. Neither temporal nor pronominal anaphora is adequately accounted for in a static semantic framework like Montague Grammar or classical (Barwise and Perry 1983) situation theory, even if the static semantic theory incorporates some contextual sensitivity such as that suggested by Kaplan (1977). The key novelty in dynamic semantics that is brought out by the characterization theorem is that in dynamic semantics it matters how information is introduced into the discourse—viz. what variable or discourse referent this information is predicated of. These discourse referents are now understood to be a fundamental part of information states and discourse contexts. Is this really part of the content? Well, they might be but it seems more intuitive to think of them as part of how information is *packaged*. Dynamic semantics invites us to make a distinction between information content and information packaging.

We can represent information packaging within possible worlds semantics at a cost. We must replace the notion of information as represented by sets of worlds with a set of n-tuples consisting of a world and a sequence of objects—one object, roughly, for each indefinite introduced into the discourse (this needs to be refined when indefinites occur within the scope of other logical operators and quantifiers). But this is a technical trick without much philosophical substance and it obscures the central point of dynamic semantics: the information conveyed by a text is more than just truth conditional content (a set of possible worlds) it also involves some sort of "information packaging" which tells us how to understand a variety of linguistic anaphoric phenomena.[6] A representational structure giving us a notion of information packaging is needed here. For dynamic semantics this information packaging is quite minimal— we can make do with just variables.

Dynamic semantics and its limited view of information packaging does not do justice, however, to the complex interaction of pragmatic and semantic factors in discourse. This leads to incorrect predictions about the way anaphors, both temporal and pronominal,are treated in DRT. In effect the notion of information packaging in DRT is too weak to support its analysis of the phenomena; it tries to do too much with truth conditions and not

[6] Stalnaker's critical evaluation in his 1996 SALT talk of dynamic semantics attempts to defend classical possible worlds semantics. But he does so by packing the representational component, information packaging, into the possible worlds themselves. Further, because he doesn't really propose a theory of information packaging, his account isn't doesn't offer an alternative to the view developed here.

enough with information packaging.

4.1 *Temporal Anaphora*

A clear example of where DRT goes wrong is with its analysis of temporal anaphora. DRT is right to make the contributions of tense depend upon the discourse context. But DRT attempts to make the temporal structure of a text almost completely dependent on the tense forms used in the text. In most natural languages, however, the temporal structure of the events introduced in a text is underdetermined by the sequence of tense forms. In particular, the rule above, which is a consequence of the DRS construction procedure, is false for French—a point of which some of the earliest workers on tense in DRT were aware—or English. Consider the following examples (from Lascarides and Asher 1993):

(5) John entered the room. Fred greeted him.

(6) John fell. Fred pushed him.

These two discourses employ the same sequence of tense forms, yet they suggest different temporal structures. DR-theorists have been forced to revise the construction procedure and to abandon the view that the tense forms and the order of the sentences in a discourse alone determine temporal structure. This conclusion follows not only from an examination of the English simple past but also from a careful look at the data concerning the French plus que parfait (Bras and Asher 1994) and the English pluperfect (Lascarides and Asher 1993b).

One might ask, what in combination with tense sequences determines the temporal structure? One proposal is that a more developed view of discourse structure determines tense structure. Originally suggested by the computer scientist Jerry Hobbs, this thesis has been worked out in the context of a formal discourse semantics by Asher and Lascarides (Lascarides and Asher 1993a, 1993b) using a more elaborate analysis of information packaging than that found in standard DRT. Specifically, it is the rhetorical connection between the propositions that often supplies the information needed to determine temporal structure.

4.2 *The Contextual Treatment of Definites*

There have been two treatments of definite descriptions vying for contention since the times of Russell and Strawson. One is Russell's famous 1905 analysis of definites in terms of first order quantifiers; the ϕ ψ's is to be translated as $\exists x(\forall y(\phi(y) \leftrightarrow y = x) \wedge \psi(x))$. Russell's analysis works well enough for the example he discusses at length,*the present King of France is not bald*; but there are other uses of definite descriptions such as (7) that render his analysis very problematic:

(7) If I invite a celebrity, the celebrity never comes.

Such uses of definites have motivated an entirely different sort of analysis from Russell's that begins with Strawson's paper of 1950 and that has proved popular amongst linguists. Most of those who have adopted this position have argued that definite descriptions generate presuppositions of familiarity; the individual that is the denotation of the definite must be familiar to the speaker's audience or already mentioned in the text. More generally, such presuppositions must be understood as holding in the context in which the sentence generating the presuppositions is to be interpreted. In the linguistic literature, this notion of holding is glossed with the help of two other notions: binding and accommodation. The presupposition of familiarity is said to be bound if the discourse referent or variable introduced by the definite is identified with a discourse referent or variable introduced by some antecedent NP. All presuppositions must either be bound or accommodated in the discourse context. When such an antecedent NP is not available, advocates of this approach say that the presupposition is accommodated by adding to the antecedent context a variable with the appropriate properties. Presupposition theorists have not said much about uniqueness— perhaps justifiably so in view of the apparent failure of any uniqueness claim or presupposition in (7). One of the success stories of dynamic semantics has been to give structure to contexts through the recursive structure of conditions so that constraints could be placed on accommodation (van der Sandt 1992); dynamic semantics made sense of the inherent dynamism in the notions of presupposition already evident in Gazdar's work in the late seventies.[7]

The presuppositional view of definites gives a nice analysis of the anaphoric behavior of the definite in (7), and it can be extended by using accommodation to handle Russell's example. However, neither the presuppositional view nor Russell's makes sense of the behavior of definites in so called "bridging" examples, where an antecedent for the definite is not explicitly given in the text but constructed from other information sources. Here are some examples.

(8) Mary moved from Brixton to St. John's Wood. The rent was less expensive

[7] For details see Beaver (1994, 1996).

(9) I met two interesting people last night at a party. The woman was a member of Clinton's Cabinet.

It is easy to see why the presupposition theory of definites (let alone the Russellian theory) fails to account for these examples. Consider for a moment (9). According to the approach to presuppositions mentioned above, the woman must be bound to some woman already introduced in the text, or its presupposition of familiarity must be accommodated. Since no woman has been explicitly mentioned prior to the discourse, one is added to the context, which gives the wrong the truth conditions.

Several proposals have been devised to supplement the presupposition theory to get the right sort of anaphoric connections for the definites in (8)-(9) — in particular the addition of lexical information or world-knowledge (Bos et al. 1995). It is possible to get the right anaphoric connection for the woman in (9) if we develop extended rules for binding that exploit the information that *woman* is a subtype of the noun *people* in the antecedent noun phrase. But this sort of approach won't make sense of (8). No lexical information or nonlinguistic bit of world knowledge about the parts of London will yield the right interpretation of the definite the rent. World knowledge about London will rather predict the opposite, since Brixton is commonly acknowleged to have apartments and houses with lower rents than St. John's Wood, an interpreter relying solely on world knoweldge or lexical information would conclude that the rent referred to the rent of a dwelling in Brixton. But this is plainly not what was intended.

What sort of information is required to get the right interpretations of these definite descriptions? An important clue to what seems to be the right answer comes from investigating parallel uses of indefinites, as this example from Charniak (1983) attests:

(10) a. Jack was going to commit suicide. He got a rope on Tuesday.

b. Jack was going to commit suicide. He got the rope on Tuesday.

An interpreter is likely to construct the same link in (10a-b) between the propositions expressed by the two clauses: Jacks gets the rope he uses to commit the suicide on Tuesday. It is the rhetorical relation between the two propositions that enables the interpreter to resolve the definite in the appropriate way.

Information about rhetorical structure also governs the interpretation of recalcitrant examples like (8). Our approach has a much more liberal notion of binding than the traditional one; identity is only one of many ways in which two discourse referents can be linked. In (8) the second sentence explains why the agent moved, and it this relation of Explanation that guides the binding of the definite *the rent*. The rent's being less expensive in St. John's would explain why the agent moved, and so the inference to an Explanatory connection between the two clauses ensures that the discourse referent introduced by *the rent* to is linked to the disocurse referent introduced by *St. John's Wood*. The discourse structure in (9) is somewhat different; there the second sentence serves as an elaboration of the event describedin the first. The fact that we have an elaboration leads us to bind the presupposition in a particular way—viz., the woman is identified as one of the people the speaker met last night. Such connections are not part of the compositional semantics of such discourses even on dynamic semantics's conception of such. They are part of information packaging. And interestingly, these rhetorical connections between propositions in a text play an important role in interpreting not only temporal anaphora but definite descriptions.[8]

4.3 Lexical Ambiguity

Few compositional theories of meaning tackle the problem of lexical ambiguity, since they are designed to articulate principles of how word meanings combine together. But lexical ambiguity is also a matter for compositional semantics and discourse interpretation, since it is usually the case that combining words with ambiguous meanings leads to a decrease in ambiguity rather than an increase in ambiguity. Sometimes syntactic factors or semantic requirements like selectional restrictions reduce ambiguity, but in many cases the resolution of semantic ambiguities depends on other pragmatic factors — in particular, ambiguity resolution often depends on the discourse structure of the discourse context, as argued in Asher and Lascarides (1995). Consider the following example from that paper:

(11) The judge demanded to know where the defendant was. The barrister apologised and said that he was drinking across the street. The court bailiff found him asleep beneath the bar.

The word *bar* in the last sentence is ambiguous even when taken as a noun. But it is not ambiguous in this discourse. Standard disambiguation techniques in AI try to use word senses in the same clause to disambiguate other words going for a most probable interpretation (different word senses being associated with different frequencies) in a particular domain. But here this would net the wrong result, since this would predict that bar is being used

[8]This approach to definite descriptions is worked out in detail in Asher and Lascarides (in press).

in its courtroom sense. But if we reconstruct what is going in the discourse and exploit the way the various sentences are related to each other in disambiguation, we get a quite different story— one that is recounted in Asher and Lascarides (1995). Very roughly, the text is a narrative; the proposition expressed by each sentence is related by the discourse relation of Narration to the proposition expressed by the previous sentence. Narration imposes strong spatio-temporal constraints on the actors and events involved, so that once we expect that the defendant is drinking at a pub, we assume as a default that he stays there unless the narrator indicates otherwise. So we interpret the last sentence such that the bailiff finds the defendant in the drinking establishment not in the court slumped underneath the bar there. Once again discourse structure seems to play an important role in discourse interpretation, and it is a role that cannot be captured within dynamic semantics as it stands.

4.4 Propositional and Concept Anaphora

If pronominal anaphoric reference to indefinites identifies a need for discourse referents in the analysis of information, then anaphoric reference to chunks or segments of text suggests other representational commitments. Consider the following text:

(12) One plaintiff was passed over for promotion three times. Another didn't get a raise for five years. A third plaintiff was given a lower wage compared to males who were doing the same work. But the jury didn't believe this (any of this).

There appear to be two possible antecedents (depending to some degree on how this is stressed) in the fourth sentence of (8)—the proposition expressed by the discourse as a whole and the proposition expressed by the last sentence. But in dynamic semantics neither proposition as it stands is a possible antecedent. It was such references to abstract objects that motivated Asher (1993) to develop a non-trivial extension to Kamp's DRT in which a richer conception of information packaging segments the information given in a discourse into bits that are supported by the data such as in (8).

Similarly, concept anaphora can be subject to the effects of discourse structure. One form of concept anaphora is verb phrase ellipsis, exemplified in the following examples.

(13) a. John said that Mary hit him. But Sam did.

b. John said that Mary hit him. Sam did too.

Aided by prosodic stress which gives certain important clues about the discourse context, speakers almost uniformly prefer the embedded verb phrase meaning to go in for the ellided VP in (13a)—i.e. $\lambda x(x hit\ y)$, while in (13b) they prefer the verb phrase meaning derived from the main VP—$\lambda x(x said\ that\ Mary\ hit\ x)$. A theory that does not take account of the discourse structure generated by particles like *but* or *too* will not be able to make sense of these preferred interpretations.

5 Beyond Dynamic Semantics: SDRT_0

A fuller theory of discourse interpretation and a theory of the pragmatics/semantics interface incorporates pragm-atics and other information sources besides just the compositional semantics exploited in DRT. In particular, such a theory must build, I have argued, interpretation from discourse structure. There are many ways of incorporating discourse structure within dynamic semantics, but the simplest is the one which remains squarely within the confines of the first order theory discussed above. The relevant discourse structure for temporal interpretation can be formalized by generalizing the idea of eventuality discourse referents introduced in section 2. In this extension of the original DRT fragment, I introduce speech act discourse referents, π_0, π_1, π_2, etc. The signature of the fragment adds two sorts of conditions to those defined above: if π is a speech act discourse referent and K a DRS, then $\pi : K \in$ *Conditions*. Discourse Structure is encoded by relation symbols on speech act discourse referents; e.g., $Elaboration(\pi_1, \pi_2)$ signifies that the speech act π_2 elaborates what was said in π_1. We extend the model theory of DRT_1 to handle conditions of the form $\pi : K$ by relativizing all conditions in K to values of π in a recursive fashion. I assume below that the interpretation of a predicate of the DRS language is relativized to M and a sequence α of the values speech act discourse referents. This relativization is intended to capture, albeit imperfectly, the intuition that the information in K pertains to the speech act discourse referent that characterizes it. We must, however, reinterpret DRSs so each condition defines a transition predicate; in effect $K^\frown\gamma$ is now the sequence $K;\gamma$. We can then ensure that variables declared in K but also occurring in K' (due to anaphoric equations) are properly bound.

Definition 2:

- $f P_M^\alpha(U, O) g$ *iff* $f \subseteq g$ *and* $dom(g) = dom(f) \cup U$

- $f P_M^\alpha(R(x_1, \cdots, x_n)) g$ *iff* $R_M^\alpha(f(x_1), \cdots, f(x_n)) \wedge f = g$

- $f P^{\alpha}_{M}(K^{\frown}\gamma) g$ *iff* $\exists h (f P^{\alpha}_{M}(K) h \wedge h P^{\alpha}_{M}(\gamma) g)$

- $f P^{\alpha}_{M}(\neg K) h$ *iff* $f = h \wedge \neg\exists g\, f\, P^{\alpha}_{M}(K) g$

- $f P^{\alpha}_{M}(K \Rightarrow K') k$ *iff* $(f = k \wedge \forall g\, (f P^{\alpha}_{M}(K) g \rightarrow \exists\, h\, g P^{\alpha}_{M}(K') h))$

- $f P^{\alpha}_{M}(K \vee K') k$ *iff* $(f = k \wedge \exists\, g\, f P^{\alpha}_{M}(K) g \vee \exists h\, f P^{\alpha}_{M}(K') h)$

- $f P^{\alpha}_{M}(\pi : K) g$ *iff* $f P^{\alpha^{\frown} f(\pi)}_{M}(K) g$

The last clause of the recursive definition above shows how the relativization of the interpretation of conditions in a DRS K is built up recursively when the latter itself occurs in a condition of the form $\pi : K$. Notice also how the last clause ensures that the output assignments from labelled constituents in which variables are declared will make those assignments available to the interpretation of subsequent labelled constituents. Otherwise, the definition is similar to the one for classic DRT. One can easily define the "lifting" of P above on the informationally empty discourse context σ_0; it is simply the same (and I'll use the same notation P) as given in section 2 for DRT. When the contexts are not informationally empty, matters become more complicated, as we shall soon see.

Let us call SDRT_0 the theory which delivers those discourse representations, which I'll call SDRSs, whose syntax and semantics is as described above. The language of SDRT_0 is first order. For the Characterization Lemma for DRT applies equally well to SDRT_0. Recall that the characterization lemma comes in two parts; the first claims that every DRS corresponds to a first order formula and the second that every first order formula corresponds to a DRS. To show that a reformulation of the Characterization Lemma holds for SDRT_0, we need only to show that every SDRS in this fragment corresponds to a first order formula. To show this, we have to specify in the induction on the complexity of SDRSs what formula corresponds to the new condition of the form $\pi : K$. We will assume that the model here has parametrized interpretations corresponding to the function V^{α} and that each predicate is parametrized for one argument, a sequence of speech acts. By the inductive hypothesis, K has a parametrized characteristic formula $\chi_K(x)$, where x is the sequence of variables corresponding to the speech acts. So the corresponding formula for $\pi : K$ is just like $\chi_K(x^{\frown}u_{\pi})$, where u_{π} is a new free variable appended to the sequence of speech acts that constitute an argument of each atomic formula in χ_K. Thus, we have established:

Characterization Lemma for SDRT_0:

- For every first order formula χ with a set of free variables U, there is an SDRS (U, C) such that $\mathcal{P}(U, C)[\sigma_0] = \{(M, f) : Dom(f) = U \textit{ and } M \models \chi[f]\}$
- Every SDRS of SDRT_0 (U, C) has a characteristic formula χ where U is the set of free variables in χ and $\mathcal{P}(U, C)[0] = \{(M, f) : Dom(f) = U \textit{ and } M \models \chi[f]\}$

SDRT_0 can be used to address many of the issues that made us dissatisfied with dynamic semantics proper—the account of definites and presupposition, the resolution of many lexical ambiguities, and some of the problems with temporal anaphora. To address these issues, however, SDRT_0 requires a background theory in which the semantic consequences of discourse relations are specified. These consequences include not only temporal effects, of which I will give an example below, but also spatio-temporal effects, effects on lexical choice and effects on the choice of binding relation and antecedent for presupposition.[9]

This background theory is most naturally expressed in the SDRT_0 language.[10] To economize on space, however, I shall use the first order translations here to give an indication of the axioms in the background theory. For example, Narration demands a consistency and coherence which many have observed and which is axiomatized in SDRT below: in words it says that if two speech acts form a narrative sequence, then the eventualities described by these speech acts must cohere together at least in the sense that the post-state of the first be consistent with the pre-state of the second (for more discussion see Asher 1996, Asher et al. 1995). I designate the main eventuality introduced in a DRS K (since we have DRSs for each clause this will be the eventuality discourse referent introduced by the main verb) by e in the characteristic formula of $\pi : K$, the prestate of an eventuality e by pre(e), the poststate of an eventuality e by post(e), the relation of temporal overlap by O_t, the relation of temporal precedence by $\prec$ and the relation of temporal inclusion by $\subseteq_t$:

$$(Narration(\pi_1, \pi_2) \wedge \chi_{\pi_1:K_1}(e_1) \wedge \chi_{\pi_1:K_2}(e_2)) \rightarrow O_t(post(e_1), pre(e_2)$$

Other uncontroversial axioms about eventualities and their pre-states and

[9]For details concerning these other semantic consequences, see Asher and Lascarides 1995, forthcoming, and Asher et al. 1995. This background theory elaborated in these articles defines semantic consequences of eight of the most prominent discourse relations found in expository text.

[10]In Lascarides and Asher 1993a, this background theory was made part of another component of reasoning—what I call below the "glue logic". But the consequences of discourse structure should be expressed as DRS or SDRS conditions, since the SDRS language is what gives the content of the discourse.

post-states ensure that we can conclude the following fact:

Fact 1 $(\text{Narration}(\pi_1, \pi_2) \wedge \chi_{\pi_1:K_1}(e_1) \wedge \chi_{\pi_1:K_2}(e_2)) \rightarrow e_1 \prec e_2$

There are other axioms capturing the temporal consequences of other discourse relations. Here are a few examples.

$Elaboration(\pi_1, \pi_2) \rightarrow e_2 \subseteq_t e_1$
$Explanation(\pi_1, \pi_2) \rightarrow \neg e_2 \prec e_1$
$(Explanation(\pi_1, \pi_2) \wedge (\neg(State(e_2) \vee Process(e_2))) \rightarrow e_2 \prec e_1$

This is not an exhaustive list of the axioms for the temporal consequences of discourse structure, let alone a list of the axioms needed for lexical disambiguation or the analysis of presuppositions. But these axioms do give an idea of how one could use these temporal order axioms and other axioms to compute the contribution of discourse relations to the interpretation of constituents, supplementing the compositional semantics. We can derive formulas about the temporal order of eventualities mentioned in a text from the SDRT_0 background theory. This enables us to show that SDRT_0 is a non-conservative extension of DRT_1 in the following sense. Recall discourse (4) above:[11]

Fact 2 *There is a formula ϕ in the language of DRT_1 such that* $(4) \not\vdash_1 \phi$ *but* $(4) \vdash_2 \phi$.

5.1 CCPs of new information in $SDRT_0$

There are again two conceptions of CCP for an unambiguous sentence S in SDRT_0. The representational conception of the CCP of S is a relation between SDRSs—in particular between an SDRS of SDRT_0 representing the discourse context, an SDRS derived from S, which I'll assume is just $\pi_S : K_S$ where K_S is the DRS produced from S using the DRS construction procedure, and an "output" SDRS.[12] Note that each piece of new information to be added to the contextual SDRS introduces its own speech act discourse referent. But otherwise SDRT_0 conceptions are similar in appearance to those in DRT. The model theoretic conception of the CCP of an unambiguous sentence too resembles its DRT counterpart; it is a relation between contexts or sets model embedding function pairs, just as was the

[11] Of course a more complex DRS constrution procedure might also net us the appropriate temporal information for (4) but as we saw in section 4, there is ample evidence to suggest that a purely semantic construction procedure will not produce the right results.

[12] This is a simplification, since discourse structure can also occur within a single sentence. We'll see how to rid ourselves of this assumption shortly.

case with DRT_0. What changes dramatically once we move to $SDRT_0$ is the complexity of computing the change wrought in the discourse context by "adding" the information contained in S. To compute the CCP of ϕ relative to a discourse context τ, one must compute a discourse relation between π and some available "attachment" point, which is a subset of the speech act discourse referents in the discourse context.

But in order to do this, we must change our conception of what a discourse context is. It can no longer simply be an information state, for such a state does not provide us with the information by means of which to do the computation of attachment.

One approach defines the CCP of a formula to be the relation that, given a discourse context allows *any* possible attachment by means of any discourse relation to make up an allowable CCP transition. That is, the CCP of ϕ relative to a discourse context τ is that relation that yields $\tau^\frown\phi^\frown R(\pi_1, \pi_2)$, where π_2 is the speech act discourse referent of ϕ and π_1 is some speech act discourse referent in τ and R some discourse relation symbol.

This proposal, however, would not enable us to make *any* predictions about the temporal structure of a discourse such as (6), let alone the meanings of definite descriptions such as in (8). The conception of CCP needs to be more restrictive. A natural alternative is to treat the construction of the appropriate SDRS in which the relevant discourse relation attaches the new information to the given context as a black box. That is, the CCP of ϕ relative to a discourse context τ is that relation that yields $\tau^\frown\phi^\frown R(\pi_1, \pi_2)$, assuming that $\pi_2 : \phi$ and that π_1 is some "appropriate" (in a sense explicitly defined only within the black box) speech act discourse referent in τ and R the "appropriate" discourse relation used to attach the information in ϕ. While this black box approach might seem a little bit like "cheating," it is customary in semantics to argue in this way. Such an approach has the virtue that it allows us to separate out processing questions from the logical character of the final representation. This black box is in effect an oracle—call it O.

What is the logical status of $SDRT_0$ defined transitions relative to this oracle O for appropriate SDRSs? Given that the characterization lemma still holds for the representations produced by $SDRT_0$ (they are in effect just DRS's), it is a straightforward matter to adapt theorem 1 to such a conception of CCP, as labelled by SDRSs defined within $SDRT_0$. The CCP transitions are more restrictive, allowing fewer model assignment pairs to be related in a transition. Further, the underlying logical theory of $SDRT_0$ is more than just first order logic, since it contains non-logical axioms about the semantic consequences of discourse structure—e.g., the temporal axioms discussed in the previous subsection. So to distinguish the absurd formulas relative to this underlying theory from those of section 3, I'll call the the set of absurd formulas for SDRT_0, $\Phi^0_\perp$.

Letting Acc be as in the preamble to theorem 1 of section 2, each SDRS ϕ labels a transition on states $\mathcal{P}^0(\phi)$; these transitions are distinct from the simpler transitions we appealed to for DRT, because they are engendered by SDRT_0 conceptions of transitions and discourse contexts). Clearly $\mathcal{P}^0$ transitions are distinct from those transitions induced simply by DRSs; they are more restrictive. But as before we may define:
$\sigma^0_\phi = [\mathcal{P}(\phi)](\sigma_0)$.
And again we may suppose that $\mathcal{P}^0(\phi^\frown\psi) = \mathcal{P}^0(\phi) \circ \mathcal{P}^0(\psi)$. But our notion of $\underline{\leftrightarrow}^0$ is more complex because of the complexity of the update relation in SDRT_0, which outputs a set of SDRSs—one for each attachment site.
$\phi \underline{\leftrightarrow}^0 \psi$ iff $\forall \phi'$:

- $\forall \chi \exists \chi'$:
    1. $(T^0(\phi, \phi', \chi)$ is defined iff $T^0(\psi, \phi', \chi')$ is defined, and
    2. $T^0(\phi, \phi', \chi) \in \Phi^0_\perp$ iff $T^0(\psi, \phi', \chi) \in \Phi^0_\perp$

.

- $\forall \chi' \exists \chi$:
    1. $(T^0(\phi, \phi', \chi)$ is defined iff $T^0(\psi, \phi', \chi')$ is defined, and
    2. $T^0(\phi, \phi', \chi) \in \Phi^0_\perp$ iff $T^0(\psi, \phi', \chi) \in \Phi^0_\perp$

Theorem 5.1 $\phi \underline{\leftrightarrow}_0 \psi$ iff $\sigma^0_\phi = \sigma^0_\psi$; *and further,* $\underline{\leftrightarrow}_0$ *and* $\Phi^0_\perp$ *are r.e. relative to O.*

Theorem 2 is a straightforward adaptation of theorem 1 to the basic SDRT fragment and its underlying logic. Note, however, that the theorem relies on essentially an oracle for giving us the "appropriate" output SDRS. That is, our transition predicate on SDRSs relies on this oracle. Can we do any better?

To do better requires a method for calculating the more restrictive and appropriate CCP notion given an input SDRS (our representation of the given discourse context) and some new information ϕ that we also represent as an SDRS of SDRT_0. To calculate this more restrictive CCP notion, we have to: (1) put constraints on which speech act discourse referents may act as attachment points in the antecedently given discourse context, and (2) articulate mechanisms for constraining what are the admissible discourse relations by means of which we can bind the new information to the given discourse structure.

Without going too much into the gory details of how SDRS's are constructed, we can still specify the CCP of a formula SDRT_0 in more detail.

I understand the constraints and mechanisms needed to compute $SDRT_0$'s restrictive CCP notion as forming a particular logic G, a "glue" logic, for deducing a new SDRS from a contextually given SDRS and an SDRS representing new information. As SDRT has pretensions to being a computational semantic theory (see Lascarides and Asher 1993a), G should be a logic that is at least decidable and preferably lower in complexity than NP-complete. So G must be distinguished from the underlying logic of $SDRT_0$, which is at least as strong as first order logic and hence undecidable.

In SDRT, the glue logic G is a nonmonotonic logic. From a computational point of view, this adds to the complexity of G deductions, but nonmonotonicity seems unavoidable when trying to compute discourse structure. While we sometimes have sufficient information in the message itself to deduce what rhetorical function the speaker intended to have this new information serve, in many cases we do not and we must make a best guess as to what the discourse relation is. Recall, for instance, the "push-fall" example (6) above on which it seems natural to conclude that the information conveyed by the second sentence serves as an explanation for what happened in the first. But now consider (14):

(14) John fell. Fred pushed him. Unable to stop himself, John slid off the edge of the cliff to his death.

In (14) we do not infer that the proposition that Fred pushed John explains why John fell. Rather, we infer that the second sentence introduces an event that is subsequent to the first. The inference to Explanation in (6) is here defeated by further information—a trademark of nonmonotinicity.

To build SDRSs, the glue logic G must exploit a variety of information sources-lexical information and information about the structure of an SDRS principally, but it may also make use of nonlinguistic information or world knowledge. As G should be a computationally tractable logic, the language of G, L_G, should have a tractable semantics. To date it has been sufficient to make the language L_G a quantifier free fragment of a first order language augmented by a weak conditional operator $>$, which formalizes generic or defeasible rules of interpretation ($A > B$ means "if A then normally B"). This language has the following syntax with formulas defined recursively from predicates Ψ and constants α.

- α ::= Speech Act Discourse Referents of SDRT
- 1 -place predicates Ψ^1 ::= $\{\phi : \phi$ *is an SDRS condition* $\} \cup \{$*Event-proposition*, *Stative-proposition*$\}$
- 2-place predicates $\Psi^2; = \{\langle ., . \rangle\} \cup \{$ *D: D is a discourse relation symbol* $\}$

- $\Phi \ ::= \ \Psi^n(\alpha_1, \cdots, \alpha_n)|\neg\Phi|\Phi_1 > \Phi_2|\Phi \wedge \Phi_2$

Predicates like *Event-proposition* and *State-proposition* are needed to capture aspectual information in SDRSs for the purpose of calculating temporal structure. But these are in effect SDRS conditions—they are just the typing conditions for certain types of discourse referents. We define the semantics of L_G relative to a world $w \in W$ and a selection function $\star : Pow(W)XW \rightarrow Pow(W)$, on which we can place certain constraints to get a particular nonmonotonic logic.[13] The modal semantics of L_G is itself rather unexceptional:

Semantics for L_G

- The usual rules for the truth of atomic and Boolean combinations of formulas apply.
- $[\![\phi > \psi]\!]_w = 1$ *iff* $* (\parallel \phi \parallel, w) \subseteq \parallel \psi \parallel$

The monotonic notion of validity in G is easily axiomatized The definition of the notion of nonmonotonic consequence relation $\mid\!\sim$ is more involved. Essentially, the idea is to turn $>$ into $\rightarrow$ whenever this is consistent. There are several ways of working out this intuition technically—the most elegant being that found in Morreau (1995). $\mid\!\sim$ gives rise to a proof theoretic equivalent.[14]

To turn G into a logic for building SDRSs, we need to add axioms for inferring discourse relations. These axioms typically exploit $>$ and the presence of certain SDRS conditions in the given SDRS (representing the discourse context) and in the SDRS representing the new information to be integrated into the context; details can be found in a number of places (e.g., Lascarides and Asher 1993a, Asher et al. 1995). But roughly they are of the following form:

$$(< \pi_1, \pi_2 > \wedge \phi(\pi_1) \wedge \psi(\pi_2)) > D(\pi_1, \pi_2)$$

which is intended to say that if π_1 and π_2 are to be related to each other in a discourse context, and conditions ϕ obtain in the SDRS characterizing π_1 and conditions ψ obtain in the SDRS characterizing π_2, then normally discourse relation D holds between π_1 and π_2.[15]

[13] For details, see Asher and Morreau (1991) or Morreau (1995).

[14] Again for details see Asher (1995) or Morreau (1995).

[15] I have changed and simplified the syntax and semantics of L_G somewhat from that suggested but not explicitly formalized, say, in Lascarides and Asher (1993a) and other works.

What we would like to say is that axioms allow us to conclude that certain discourse relations hold by default, when the consequences of their holding are consistent with what compositional and lexical semantics yields. But one cannot do that without rendering $\mid\!\sim_G$ undecidable because this would require the testing of the consistency of arbitrary first order conditions. Instead G is allowed access to certain information about conditions–including some of the consequences of discourse structure determined in the SDRT$_0$ background theory. This information is encoded in G as *coherence constraints.* These mimic the information encoded in the background theory of SDRT to a limited extent; for instance, we have:

- $Narration(\pi_1, \pi_2) \rightarrow [e_1 \prec e_2](\pi_2)$
- $[e_1 \prec e_2](\pi_2) \rightarrow \neg[O_t(e_2, e_1)](\pi_2)$

In words the first axiom says that if Narration holds then a certain condition must be true of π_2—namely that the condition $e_1 \prec e_2$ holds of π_2 or more particularly is a condition in the SDRS that π_2 characterizes. The second axiom is an example of how to encode information about the incompatibilities between various DRS conditions in L_G.

For a glue logic capable of generating appropriate SDRSs to analyze temporal anaphora and lexical disambiguation, we are able keep G at a manageable level of complexity, as was shown in Lascarides and Asher (1993a).

Fact 3 $\mid\!\sim_G$ *in SDRT is decidable.*

As the gloss on the axioms for discourse relations suggests, we work within a very particular set of models for G in using G to build SDRSs. Each world of these models is an SDRS or discourse structure, and the satisfaction of atomic L_G formulae in such models is just the presence of the appropriate condition in the SDRS. For instance:

- For individual constants π, $[\![\pi]\!] = \pi$.
- If Ψ is a discourse relation symbol, then $[\![\Psi]\!]_w = \{\langle\pi, \pi'\rangle : \Psi(\pi, \pi')$ *is a condition in* $w\}$.
- If Ψ is not a discourse relation symbol and not of the form , then: $[\![\Psi]\!]_w = \{\pi' : \exists K'\Psi$ is a condition or a generalization of a condition in K' and $\pi' : K'$ is a condition in $w\}$
- $[\![<\cdot,\cdot>]\!]_w = \{<\pi_1, \pi_2>$: For some discourse relation symbol D, $D(\pi_1, \pi_2)$ is a condition in $w\}$

We can't capture these particular models with the expressive power of L_G. But we can codify this connection between G and SDRT by means of a function μ that "transfers" information from SDRSs to formulas of L_G. This function isn't part of L_G; it's the link between G and the SDRSs G is supposed to reason about. It's the link between SDRT's logic of information packaging and its logic of information content.

Now we can express the notion of an SDRT_0 transition defined relative to some unambiguous sentence S. I define the T_0 transition relation for SDRT_0, which relates K, K' and a new SDRS $K^\dagger$. Intuitively, $K^\dagger$ is an SDRS where the old and new information have been merged together. More specifically, $K^\dagger$ includes (a) the old information K, (b) the new information K' derived from S, and (c) an attachment of K' with a rhetorical relation to an available attachment point in K. The relation $T_0(K, K', K^\dagger)$ is constrained so that it can hold only if part (c) of $K^\dagger$ is computed via G. More formally, let $Avl(K)$ be the set of available attachment sites in K, and let K_π stand for the SDRS ϕ such that $\pi : \phi$ occurs in K. Further, let $Pred_\pi$ be the label of the SDRS constituent in which π is declared or (equivalently) in which a condition of the form $\pi : K$ occurs, let K_π stand for the SDRS constituent labelled by π, and finally let $\alpha[\beta/\gamma]$ be the result of replacing γ in α with β. Then the T_0 predicate for SDRT is defined as:

- **`The Update Relation`**
  $T_0(K, K', K^\dagger)$ *iff* $\exists \pi \in Avl(K)$ such that:

  1. $(\mu(K_\pi), \mu(K')) \mid\!\sim (R(\pi, \pi') \wedge \varphi)$; and
  2. $K^\dagger = K[K^+/K_{Pred(K_\pi)}]$, where:
  3. $K^+ = Update_{drt}(K_{Pred_\pi}, \langle\{\pi'\}\{\pi' : K'(\varphi), R(\pi, \pi')\}\rangle)$, where $K'(\varphi) = K'$ together with those conditions specified in φ, where φ is that information needed to satisfy the coherence constraints on R.

The resulting SDRS for the discourse incorporates the new information and a rhetorical relation R that is computed on the basis of the axioms of G together with information about the semantic content of the old and new information (i.e., $\mu(K_\pi)$ and $\mu(K')$) where the SDRS K_π is the one characterized by π. The background SDRT theory ensures that the appropriate semantic consequences of the discourse links are included in the content of the discourse. Clearly, if the set of theorems in G is decidable or even r.e., then we can eliminate the reference to the oracle O in theorem 2 and show that SDRT_0 transitions are r.e. *tout court.* This constitutes a decided improvement over theorem 2.

6 Reasons to be dissatisfied with SDRT_0

SDRT_0 allows us to accomplish a great deal with a very little extra bit of information packaging— a vanishingly small amount in fact. But the theory is unsatisfying because of its extensional character.

First, the extensional character of the state transitions defined in SDRT_0 won't do to handle anaphoric references to abstract entities. We have imposed few constraints on the interpretation of the speech acts, and importantly none that links the speech act with the information content of what was said in the speech act. Because of this and because we have not introduced propositions as objects whose content and behavior is linked to the SDRSs that describe them, the interpretation of propositional anaphoric constructions such as

(15) John got an A on his test, but Sam doesn't believe it.

is hopeless. In principle, we could identify the discourse referent introduced by *it* in the second clause with the speech act discourse referent introduced by the first clause, but there would be nothing in the semantics that determined that what Sam didn't believe was that John got an A on his test. Further, examples like (12) can't be analyzed within the theory.

A related difficulty with SDRT_0 is that we cannot really capture the meanings of conditions of the form $R(\pi_1, \pi_2)$, when the semantics of the discourse relation R involves an appeal to the content of what was said in π_1 or π_2. We can specify the first order consequences of such conditions by means of axioms—e.g., the spatial and temporal effects of discourse relations. On the other hand, an SDRS condition like *Elaboration*(π_1, π_2) should entail that the content or proposition labelled by π_2 entails that labelled by π_2. A condition like *Explanation*(π_1, π_2) also involves a relation between the propositions labelled by π_1 and π_2. But these relations cannot be expressed in SDRT_0. It is difficult to see exactly how these relations can be treated in a purely extensional first order framework for familiar reasons.

A third reason for abandoning the extensional framework of SDRT_0 for a different theory comes with dialogue. In dialogue, a speech act may have a rhetorical or discourse function that it simply cannot have in monologue or that is very rare in monologue. SDRT represents such different functions as different discourse relations; but we can also think of these as different types of speech acts. In dialogue, there are a host of discourse relations in dialogue that cannot commit one— as one is forced to do in SDRT_0—to the truth of what the participants said. For example consider a correction like the following:

(16) a. A: John distributed the copies.

b. B: No, it was Sue who distributed the copies.

Corrections are a typical type of speech act (or discourse relation) in dialogue. In corrections, it is clear that the content of the dialogue is not the relational composition of the model assignment pairs verifying each of A's and B's assertions; for that would net us an empty truth conditional content. Rather there is a complex relation between the contents of what is said in the two speech acts; the second speech act is an attempt to correct what the second speaker sees as deficiencies in the content of the first speech act.[16]

We can just represent such disagreements as given by (16) in $SDRT_0$, because the interpretations of conditions is relativized to a sequence of speech acts. If one speech act asserts p and the another not p, $SDRT_0$ represents this disagreement without inconsistency. But it also does not represent these speech acts as *disagreeing* with each other; their contents are simply incomparable. This isn't right either. To make sense of corrections and other such speech acts in dialogue, we have to be able to talk about the contents of the two speech acts. And that we cannot really do in $SDRT_0$.

One further complication is that these different contents may involve anaphoric links such as (19):

(19) a. A: John shot a man during the robbery

b. B: No he didn't shoot him.

In order to make sense of the anaphoric connections, the contents associated with speech acts that are linked by discourse relations must have linked interpretations.

[16]Mark Danburg Wyld has made a careful study of these nonveridical or "divergent" discourse relations and speech acts for his dissertation (Danburg-Wyld forthcoming). Another kind of divergent discourse relation is *Counterevidence* of which (17) is an example.

(17) a. A: Smith shot the guard at the bank.

b. B: He has witnesses that say he was out of town at the time of the robbery.

Another type of speech act that Danburg-Wyld isolates is one in which the speaker denies a discourse relation implicated to hold between two propositions by some other speaker. Here is an example:

(18) a. A: John went to jail. He embezzled pension funds.

b. B: That's not why he went to jail. He was convicted of tax fraud.

We have uncovered three reasons for going beyond an extensional theory of discourse interpretation: a theory of abstract entity anaphora, an account of the semantics of discourse relations, and a treatment of nonextensional discourse relations in dialogue. The theory in the next section addresses these concerns.

7 An Intensional Theory of Discourse Interpretation: SDRT_1

To handle propositional anaphora and VP ellipsis, we must change the signature of SDRT_0 to include conditions that represent the identification of a discourse referent with some DR-theoretic structure that represents a proposition or a property. A first try would be to introduce conditions of the form $z = K$, where K is a DRS or a lambda abstracted DRS (representing a property or verb phrase denotation). But this isn't quite right, because K is not a singular term in the syntax of SDRT_0. We need a theory internal representative of K, $\sharp K$ for each DRS. $\sharp K$ is a singular term and denotes an object in a model of SDRT_1. So we extend the signature of SDRT_0 to include such singular terms. This marks a departure from standard DRT, in which there are no singular terms at all with a constant interpretation. We also suppose discourse referents ranging over the objects denoting by singular terms. Finally, to handle dialogue or multilogue, we will replace conditions of the form $\pi : K$ with $\pi : (x, \sharp K)$, where x is a discourse referent representing the speaker in the dialogue whose speech act π is.

As the purpose of introducing singular terms of the form $\sharp K$ was to have a way of referring to contents in SDRT, so this must be reflected in the interpretation of these terms. The idea that first comes to mind (examined in some detail for DRT in Frank (1997) and in recent unpublished work by Kamp) is to exploit the correctness definition and to assign these terms sets of world embedding function pairs, having fixed a model. This amounts to making *dynamic propositions* entities in the models. Because these terms may share discourse referents with the context in which they occur and these shared discourse referents must be assigned the same value, the interpretation of such terms will be sensitive to the discourse context in which they are to be interpreted. So we will assign to $\sharp K$ a discourse context—i.e., a set of world assignment pairs relative to a given model and assignment (elements of a discourse context). The set of world assignment pairs denoted by $\sharp K$ is determined relative in distributive fashion, in keeping with the tradition of dynamic semantics.[17] Since assignments are always finite

[17]The interpretation of conditions involving $\sharp K$ will also change so as to exploit the

and we can assign an order to the discourse referents in their domain, we can think of the assignments in these pairs as finite sequences of objects in the domain of the model. Nevertheless, this requires some minimal amount of set theory in our model and thus marks a complication in the logical foundations of discourse interpretation. On the other hand, we no longer need to relativize the interpretation of conditions to labels for SDRS constituents to interpret conditions of the form $\pi : \sharp K$, since we will have an interpretation for $\sharp K$ as a whole. So our satisfaction definition below will have one fewer parameter than the previous one. Further, we can return to treating conditions as "tests" using the interpretation function V.

To carry out these changes, we will pass from an extensional semantics to an intensional one. Our models will contain not only a domain of individuals and interpretations of the primitive nonlogical relation symbols but also a set of worlds. The interpretation of primitive relation symbols will be a function from worlds to extensions. The interpretation of $\sharp K$, however, will also necessitate the presence of certain sets in the models—sets of pairs of worlds and and assignment functions.

A potential source of problems is that we want the interpretation of $\sharp K$ to be sensitive to the assignments made to discourse referents in the SDRS in which $\sharp K$ occurs as a term, as well as to the assignments made to discourse referents in other terms $\sharp K'$ which are discourse related to $\sharp K$ and which hence may support anaphoric connections to discourse referents in $\sharp K$ as in 19. This generates two problems. The first is that the interpretation of $\sharp K$, because it relies on the state transition $P_M(K)$, must be defined relative to an input world assignment pair where the assignment does not assign objects to all the discourse referents that occur as terms in conditions of K *but* are not declared in K either (since they are declared in the universes of other K' where $\sharp K$ and $\sharp K'$ are discourse related). So we must modify the interpretation of $\sharp K$ and K to deal with such "improper" DRSs or SDRSs. The interpretation of $\sharp K$ will in fact be an interpretation of $\sharp K$ paired with the background assignment.

The second problem is that, on pain of violating well-foundedness, we don't want the assignments that are part of the interpretation of $\sharp K$ to include the assignment made to $\sharp K$. To this end, I define for any discourse referent x occuring as a term in the conditions of K but not in the universe of K:

- $f_K^- = \{< x, f(x) >: x \text{occurs in K but not in} U_K\}$

---

meaning of $\sharp K$, as will the conditions involving speech act discourse referents if these are definable in terms of the contents associated with the speech acts. I won't go into these definitions here, but for an attempt in this area see Asher 1993. Many but not all the discourse relations can be given definitions in terms of the contents associated with their arguments.

Because every SDRS is well-founded, by exploiting f^- rather than f, we will ensure that our semantics is well-founded too:

Definition 3:

- $(w, f)P_M(U, O)(w', g)$ *iff* $w = w' \wedge f \subseteq g \wedge \text{dom}(g) = \text{dom}(f) \cup U$.

- $(w, f) \in V_M(R(x_1, \cdots, x_n))$ *iff* $R_{w,M}(f(x_1), \cdots, f(x_n))$

- $(w, f) \in V_M(x = \sharp K)$ *iff* $f(x) := V_M(\sharp K, f)$

- $(w, f)P_M(K^\frown\gamma)(w', g)$ *iff* $(w, f)P_M(K)(w', g) \wedge (w', g) \in V_M(\gamma)$

- $(w, f) \in V_M(\neg K)$ *iff* $\neg\exists h\, (w, f)P_M(K)(w, h)$

- $(w, f) \in V_M(K \Rightarrow K')$ *iff* $\forall k\, ((w, f)P_M(K)(w, k) \rightarrow \exists\, h\, (w, k)P_M(K')(w, h))$

- $(w, f) \in V_M(K \vee K')$ *iff* $\exists\, g\, (w, f)P_M(K)(w, g) \vee \exists\, h(w, f)P_M(K')(w, h)$

- $(w, f) \in V_M(\pi : x, \sharp K)$ *iff* $\langle f(\pi), f(x), V_M(\sharp K)\rangle \in Say_{w,M}$

- $V_M(\sharp K, f) = \{(w', h) : \exists g(f_K^- \subseteq g \wedge (w', g)P_M(K)(w', h) \wedge \text{dom}(g) = \text{dom}(f_K^-) \cup \{x : x$ occurs in a condition of K but not in $U_K\})\}$

As the satisfaction definition stands, the interpretation of terms of the form $\sharp K$ is still not sufficiently constrained. As I mentioned above, anaphoric relations may obtain between components of an SDRS that are discourse related to each other. This means that two complex singular terms $\sharp K$ and $\sharp K'$ may share discourse referents. To have the anaphoric links make sense, we must give the same assignments to the shared discourse referents. The constraint (on models) that the contents associated with any two discourse linked speech acts be *equipollent* accomplishes this. What the two subclauses in the definition of equipollence do is to set up a bisimulation between the contents of the two related speech act with respect to the assignments to the common discourse referents that appear in their respective characterizing SDRSs.[18]

Definition of Equipollence: $\sharp K_1$ and $\sharp K_2$ are equipollent relative to M, written $\sharp K_1 \sim_M \sharp K_2$, iff:

- $\forall w', h(w', h) \in V_M(\sharp K_1) \rightarrow \exists g \exists w''((w'', g) \in V_M(\sharp K_2)) \wedge$
  $\forall x \in Dom(h) \cap Dom(g) h(x) = g(x)))$
- $\forall w', h((w', h) \in V_M(\sharp K_2) \rightarrow \exists g \exists w''((w'', g) \in V_M(\sharp K_1) \wedge$
  $\forall x \in Dom(h) \cap Dom(g) h(x) = g(x))$

Equipollence Constraint

- $\forall w', g((w', g) \in V_M(\pi_1 : x_1, \sharp K_1) \cap V_M(\pi_2 : x_2, \sharp K_2)$
  $\wedge R_{w',M}(g(\pi_1), g(\pi_2))) \rightarrow \sharp K_1 \sim_M \sharp K_2$

$\sim_M$ is an equivalence relation. So we can show that if $\sharp K \sim_M \sharp K'$ and $\sharp K' \sim_M \sharp K''$ that we can show that there is a bisimulation on the values of $\sharp K$ and $\sharp K''$ with respect to the discourse referents common to $\sharp K$, $\sharp K'$ and $\sharp K''$, which is what we really need to assure semantic coreference in the anaphoric equations.

The syntax and semantics of $SDRT_1$ allows us to identify pronouns referring to propositions, facts or properties with the values of $\sharp K$ or with the value of a lambda abstracted SDRS.[19] A discourse context is on this view also a dynamic proposition, consisting of a relational structure of dynamic propositions. In $SDRT_0$ we had no propositional constituents in the discourse context— only speech acts. In $SDRT_1$ dynamic propositions are full citizens.

Every SDRS of $SDRT_1$ has a translation into a higher order intensional logic as, say, developed in Gallin (1975). We have an additional type in our logic besides that of worlds, entities and truth values: the type of an assignment. To ensure that this type is well-founded, we will have to build it up inductively in the following fashion using functional types and Currying cartesian products. Let π be the type of discourse referents or stores in our type theory (as in Muskens 1996):

- $g_0 : \pi \longrightarrow e$

[18]You might think that equipollence is too strong and all that is needed is some conditional dependence–like that given by the first clause of equipollence, but this would allow us to have some assi gnments in the meaning of $\sharp K_2$ that did whatever they wished to the drefs or variables declared in K_1. To rule out this unwanted possibility, the second clause needs to be added.

[19]To do the latter, it would appear attractive to have a third store of assignments for VPs that we can pick up and use in new contexts, but I won't work out the details here.

- $g_{n+1} : \pi \to (w \to (\bigcup_{i\Box n} g_n \to t)$
- $g = \bigcup_{n \in \omega} g_n$

Let's call the higher order logic with this set of types TY_3*. Besides the basic DRT translation into first order logic we now have the translation of dynamic propositional variables as variables of type $w \to (g \to t)$, while the translation of $\sharp K$: $tr(\sharp K) = {}^\wedge tr(K)$. We then exploit this translation in the following characterization lemma, together with the "lifting" of P defined above to the empty discourse context: $\sigma_0^1 = \{(M, w, 0) : M \textit{ is an intensional model}, w \in W_M\}$.

Characterization Lemma for SDRT$_1$:

- For every higher order intensional formula χ with a set of free variables U, there is an SDRS of SDRT$_1$ (U, C) such that $\mathcal{P}^1(U, C)[\sigma_0^1] = \{(M, f) : Dom(f) = U \textit{ and } M \models \chi[f]\}$
- Every SDRS of SDRT$_1(U, C)$ has a characteristic formula χ of higher order intensional logic where U is the set of free variables in χ and $\mathcal{P}^1(U, C)[\sigma_0^1] = \{(M, f) : Dom(f) = U \textit{ and } M \models \chi[f]\}$

As in Muskens's (1996) system, we have lost in SDRT$_1$ the SDRT$_0$ equivalence between a logic with purely objectual variables and the target dynamic theory. For we now have to introduce discourse referents as a primitive type into the intensional logic to get the equivalence with SDRT.

The work of Cocchiarella (1989) has shown how to build axiomatizations of theories with stratified comprehension, and our "stratification" of propositional types in TY_3* can borrow these ideas to get an axiomatization, which together with the techniques of Fernando leads to a notion of bisimulation over transitions defined by SDRS's intensionally construed and simultaneously over the state transitions defined by their model theoretic interpretations in the same general vein as theorems 1 and 2. All the same definitions pertinent to the definition of $\underline{\leftrightarrow}^1$, the notion of transition equivalence for SDRT$_1$, and for $\Phi_\perp^1$ and accessible states carry over here. It is a simple exercise to modify the definition of the update relation in SDRT$_0$ to fit SDRT$_1$. The update relation exploits only the representational structures and the glue logic G. The syntax of the representations in SDRT$_1$ has changed from SDRT$_0$ but only in ways that are inessential to the definition of the transition relation.

A natural query is to ask whether one could eliminate the complexity of the well-founded type of embedding functions in the models. In fact it appears that one can and thereby get a much simpler intensional version of SDRT, which I'll call $SDRT_2$.[20]

The idea here is to make the assignments relevant to interpreting the intensional contexts part of the overall context and not part of the semantic value of the intensional terms. Again, we want the interpretation of $\sharp K$ to be sensitive to the assignments made to discourse referents in the SDRS in which $\sharp K$ occurs as a term, as well as to the assignments made to discourse referents in other terms $\sharp K'$ which are discourse related to $\sharp K$ and which hence may support anaphoric connections to discourse referents in $\sharp K$ as in 19. The interpretation of $\sharp K$ relies on the state transition $P_M(K)$; but it must be defined relative to an input world assignment pair where the assignment assigns objects to all the discourse referents that occur as terms in conditions of K *but* are not declared in K either (since they are declared in the universes of other K' where $\sharp K$ and $\sharp K'$ are discourse related). Further of course, these input world assignment pairs must respect the assignments made to discourse referents by the "outside" assignment. To do this, I will need to keep track of not only the outside assignment but also of previous assignments in previous intensional contexts in order to interpret conditions of the form $\sharp K$. And to do this I will redefine the notion of a context to have two assignments—one for the extensional discourse referents and one for the "intensional" discourse referents (the discourse referents that occur within terms of the form $\sharp K$). The conditions of the form $\pi : x, K$ will not then function simply as tests as in definition 3 but will actually change the assignments to the intensional contexts, in a way similar to definition 2.

Definition 4:

- $(w,f,k)P_M(U,O)(w',g,k')$ *iff* $w = w' \wedge k = k' \wedge f_w \subseteq g_w \wedge \mathrm{dom}(g,w) = \mathrm{dom}(f,w) \cup U$

- $(w,f,k)P_M(R(x_1,\cdots,x_n))(w',f',k')$ *iff* $w = w' \wedge f = f' \wedge k = k' \wedge R_{w,M}(f_w(x_1),\cdots,f_w(x_n))$

- $(w,f,k)P_M(x = \sharp K)(w',f',k')$ *iff* $w = w' \wedge f = f' \wedge k = k' \wedge f(x) := V_M(\sharp K, k, w)$

- $(w,f,k)P_M(K^\frown\gamma)(w',f',k')$ *iff* $\exists w'',f'',k''$ $((w,f,k)P_M(K)(w'',f'',k'') \wedge (w'',f'',k'')P_M(\gamma)(w',f',k'))$

[20] Something like this approach has been suggested for the treatment of conditionals by Matthew Stone (1997) and Robert van Rooy (1997).

- $(w, f, k)P_M(\neg K)(w', f', k')$ *iff* $w = w' \wedge f = f' \wedge k = k'$
  $\neg\exists h\ (w, f, k)P_M(K)(w, h, k)$

- $(w, f, k)P_M(K \Rightarrow K')(w', f', k')$ *iff* $w = w' \wedge f = f' \wedge k = k'$
  $\forall g\ ((w, f, k)P_M(K)(w, g, k) \rightarrow \exists\, h\ (w, g, k)P_M(K')(w, h, k))$

- $(w, f, k)P_M(K \vee K')(w', f', k')$ *iff* $w = w' \wedge f = f' \wedge k = k'$
  $\exists\, g\ (w, f, k)P_M(K)(w, g, k) \vee \exists\, h(w, f, k)P_M(K')(w, h, k)$

- $(w, f, k)P_M(\pi : x, \sharp K)(w', f', k')$ *iff* $w = w' \wedge f = f' \wedge$
  $\langle f(\pi), f(x), V_M(\sharp K, f, k, w)\rangle \in Say_{w,M}$
  $\wedge\ \forall w' \in V_M(\sharp K, f, k, w)(w', f \cup k, f \cup k)P_K(w', f \cup k, k')$

- $V_M(\sharp K, f, g, w) \;=\; \{w' :\ \exists k\ (w', f \cup g, f \cup g)P_M(K)(w', k, k))\}$

This way of proceeding is in some ways much closer to the original formulation of SDRT_0. We don't need a notion of equipollence because the intensional assignments are updated each time a conditon of the form $\pi : K$ is interpreted and so variables within the constituent K that are declared in other portions are interpreted appropriately. The propositions that we refer to in SDRT_2 aren't *dynamic* but rather static; all dynamic elements are located in the assignment functions, which are not part of semantic values in the interpretation of SDRT_2.

The real payoff of SDRT_2 is a correspondence with a standard intensional logic with only objectual variables. Every SDRS of SDRT_2 has a translation into higher order intensional logic as, say, developed in Gallin (1975). Besides the basic DRT translation into first order logic we now have the translation of $\sharp K$: $tr(\sharp K) = {}^\wedge tr(K)$. We then exploit this translation in the following characterization lemma, together with the "lifting" of P defined above to the empty discourse context: $\sigma_0^2 = \{(M, w, 0, 0) : M\, is\ an\ intensional\ model, w \in W_M\}$.
Characterization Lemma for SDRT_2:

- For every higher order intensional formula χ with a set of free variables U, there is an SDRS of $\text{SDRT}_2(U, C)$ such that $[\mathcal{P}^1(U, C)](\sigma_0^2) = \{(M, w, f, g) : M \models \chi[f \cup g]$, where $Dom(f) = U \wedge Dom(g) = U*$, and where $U*$ is the set of free variables x_i such that for some formula ϕ, ${}^\wedge\phi$ occurs in χ and x_i occurs in $\phi\}$

- Every SDRS of $\text{SDRT}_2(U, C)$ has a characteristic formula χ of higher order intensional logic where U is the set of free variables in χ, $U*$

is the set of free variables x_i such that for some formula ϕ, $^\wedge\phi$ occurs in χ, x_i occurs in ϕ and $\mathcal{P}^1(U,C)_{\sigma_0^2} = \{(M,f,g) : Dom(f) = U, Dom(g) = U* \text{ and } M \models \chi[f \cup g]\}$

This characterization lemma, in conjunction with the use of generalized Henkin models for higher order modal logic, leads to a notion of bisimulation over transitions defined by SDRS's intensionally construed and simultaneously over the state transitions defined by their model theoretic interpretations in the same general vein as theorems 1 and 2. Once again the definitions pertinent to the definition of $\underline{\leftrightarrow}^2$, the notion of transition equivalence for $SDRT_2$, and for $\Phi^1_\perp$ and accessible states carry over from our earlier theories or are easily modified as is the definition of the update relation.

Given that $\mid\!\sim_G$ (including the coherence constraints) is decidable, we have the following:

Theorem 7.1 $\phi\underline{\leftrightarrow}^1\psi$ iff $\sigma^2_\phi = \sigma^2_\psi$; *and further,* $\underline{\leftrightarrow}^2$ *and* $\Phi^2_\perp$ *are r.e.*

8 More Information Packaging: $SDRT_3$

$SDRT_1$ and $SDRT_2$, like $SDRT_0$, have a small amount of information packaging. The advantage of this is that it is easy to extend the correspondence between model-theoretic and syntactically or representationally defined state transitions that is found in dynamic semantics. We can in fact do alot with these theories. We can represent various forms of ellipsis and anaphora having to do with reference to abstract objects; we can study dialogue; and we can analyze the other phenomena mentioned in connection with $SDRT_0$.

But a proper theory of anaphora and ellpsis still remains outside the purview of $SDRT_{0,1,2}$. These theories of discourse interpretation do not reflect, for instance, important pragmatic constraints on anaphora. A widely noticed fact is that discourse structure governs which antecedents to pronouns and which temporal discourse referents are permissible antecedents. It is very difficult to refer with a pronoun to an object mentioned in some proposition that is not discourse related to the proposition expressed by the sentence containing the pronoun or on the "right frontier" of the discourse structure. But as a study of the satisfaction definitions 2 and 3 shows, a pronoun can refer to the value of any discourse referent introduced in the discourse, no matter how far back.

A plausible-sounding solution to our first difficulty is to make use of some sort of "down-date" of the assignment functions. But already this makes much more complex the problem of getting a correlation between representationally defined and model-theoretically defined transitions. Second, while pronominal anaphora and temporal anaphora obey an "availability" constraint imposed by discourse structure, studies of texts have shown that definite descriptions may pick up antecedents that do not meet this constraint (Asher 1993). So it looks as if downdating the assignments (removing some variables from their domains) will not work for definite descriptions. We need a richer notion of information packing within which to represent the different status of discourse referents. In particular, this richer notion must reflect aspects of the structure of an SDRS.

An additional advantage of adding structure is that we can internalize within the notion of a discourse context the G governed notion of CCP and make sense of the information transfer function as being part of our notion of a discourse context. If we want to have a clear conception of the state transition engendered by new information that comes from a variety of sources after the processing of the verbal message, then it behooves us to incorporate the additional information needed to compute the transition into our representation of the discourse context. So far we have not constructed a completely model theoretic notion of CCP for SDRT; we have merely used the relational composition of DRT and exploited the link between the model theory and the representational formulations of dynamic discourse interpretation. But this doesn't really give us a theory in which the model-theoretic notions of context are incrementally built up as new information is added. By adding to information packaging information about the structure of the SDRS along the lines of Asher (1996), we will be able to do so.

A related issue is that even in $SDRT_{1,2}$ we have a lousy semantics for attitudinal constructions, a semantics which does not distinguish between logically equivalent beliefs. One solution to the problem of logical equivalence of beliefs is to exploit the structure of the object of belief in a context sensitive way (see Asher 1986, Kamp 1990, or Perry and Crimmins, for example. If these objects of belief are just the sort of propositions that are involved in our representation of discourse (and why not in view of the intimate link between saying and believing), then this also suggests that our view of information structure should at least include the logical structure of the SDRS and maybe even something about the concepts that figure within the conditions of the SDRS. Once again the whole SDRS seems to be relevant to information packaging—and not only for pragmatic processing issues but for the semantics of certain constructions as well.

A sufficient enrichment of information packaging to address the concerns raised in the previous paragraphs requires several modifications to our conception of a discourse context. First we will want to represent explicitly the structure of an SDRS— we can do this by encoding this structure on the set of speech act discourse referents. Second, we need to say something about the structure of the contents—the SDRSs—associated with each speech act, at least enough to be able to use the "glue" logic to compute the appopriate discourse relation for attachment, given a DRS—i.e. a compositional semantic representation of the new information—together with lexical information and pragmatic principles. We now make these changes more precise on the conceptually simpler version $SDRT_2$.

To this end let a discourse context (relative to a chosen model) be a tuple consisting of a world, as before, and, in addition, a triple consisting of the set of speech act discourse referents X, a strict partial ordering R on X, a function μ from X into formulas of the "glue" logic encoding information about the conditions associated with that element of X, and an element $x^0 \in X$ designating the current speech act. μ is our information transfer function; if x is not a speech act discourse referent $\mu(x)$ is the conjunction of all the conditions in which x figures as an argument or if x is a speech act discourse referent then $\mu(x)$ is the conjunction of all the conditions in the DRS representing the content of the speech act x; in other words μ keeps a record of the actual conditions used in the discourse and files them with each associated discourse referent.

- (w, X, R, μ, x^0)

x^0 and R together tell us which speech act discourse referents may act as suitable attachment points among the set X of all speech acts in the antecedent context. Most researchers in discourse theory have made the "right frontier" of the discourse structure the area in which possible attachments of new information may be made (though for some complicating details see Asher 1993), a constraint easily representable with our new notion of discourse context.

We then define that two such contexts τ and σ stand in the CCP relation inductively similarly to the way we defined CCP transitions between simpler contexts in definitions 2 -4. As in definition 4, the world and assignment components of the discourse contexts do all the work in the definition of satisfaction of conditions. The principal change is the way information packaging gets treated— viz. in the first clause in the recursion. The input and output discourse contexts' information packaging is now much more finely structured, and so this requires several clauses to make sure that each part of the structure is modified in the appropriate fashion. Further, as the inference of the appropriate discourse connection between the new information and the context makes use of conditions in SDRSs, we must define the transition P in a somewhat different way; this recursive definition defines P for an SDRS with an arbitrary set of conditions. As in definitions 2, 3 and 4, we will suppose that $M_\tau = M_\sigma$ in any CCP transition defined by an SDRT_1 representation, and we will make the valuation V of conditions sensitive to contexts, as in definition 3. σ_0^3 represents the empty discourse context in SDRT_3, and DR is the set of discourse relations.

Definition 5:

- $(\sigma_1, f, k)P_M(U, Con)(\sigma_2, g, h)$ *iff*:
  1. $w_{\sigma_1} = w_{\sigma_2} = w$
  2. $f_w \subseteq g_w \ \wedge \ dom(g, w) = dom(f, w) \cup U$
  3. $\forall \gamma \in Con \ (\sigma_2, g, k)P_M(\gamma)(\sigma_2, g, h)$
  4. $X_{\sigma_2} = X_{\sigma_1} \cup \{\pi : \pi \in U\}$
  5. $x^0_{\sigma_2} \in \{\pi : \pi \in U\}$, if $\{\pi : \pi \in U\}$ is non-empty; $x^0_{\sigma_2} = x^0_{\sigma_1}$, otherwise; and $x^0_{\emptyset} = \emptyset$.
  6. $\mu_{\sigma_2} = \mu_{\sigma_1} \cup \{(\phi, x^0_2) : \phi \in Con\}$
  7. *either* $(\sigma_1 = \emptyset \wedge R_{\sigma_2} = R_{\sigma_1} = 0)$ *or* $\exists \pi \in X_{\sigma_1} \exists D \in \mathrm{DR}$ $(R_{\sigma_1}(\pi, x^0_{\sigma_1}) \wedge (\sigma_2, g, h)P_M(D(\pi, x^0_{\sigma_2}))(\sigma_2, g, h)$ $\wedge(\langle \pi, x^0_{\sigma_2}\rangle \wedge \mu_{\sigma_1}(\pi) \wedge \mu_{\sigma_2}(x^0_{\sigma_2}) \mid\!\sim_G D(\pi, x^0_{\sigma_2}) \wedge$ $R_{\sigma_2} = \textit{Transitive Closure}(R_{\sigma_1} \cup (\pi, x^0_{\sigma_2}))$
- $(\sigma, f, g)P_M(R(x_1, \cdots, x_n)(\sigma, f', g')$ *iff* $\wedge f = f' \wedge g = g'$ $(R_{w_\sigma, M}(f_{w_\sigma}(x_1), \cdots, f_{w_\sigma}(x_n))$
- $(\sigma, f, g)P_M(x = \sharp K)(\sigma, f', g')$ *iff* $f(x) := V_M(\sharp K, f, \sigma) \wedge$ $f = f' \wedge g = g'$
- $(\sigma, f, g)P_M(\neg K)(\sigma, f', g')$ *iff* $\neg \exists h \exists \sigma' \, (\sigma, f)P_M(K)(\sigma', h)$ $\wedge f = f' \wedge g = g'$
- $(\sigma, f, g)P_M(K \Rightarrow K')(\sigma, f', g')$ *iff* $\forall k \forall \sigma' \, ((\sigma, f)P_M(K)(\sigma', k)$ $\rightarrow \exists \, h \exists \, \sigma'' \, (\sigma', k)P_M(K')(\sigma'', h)) \wedge f = f' \wedge g = g'$
- $(\sigma, f, g)P_M(K \vee K')(\sigma, f', g')$ *iff* $\exists \, g \, \exists \, \sigma' \, (\sigma, f)P_M(K)(\sigma', g)$ $\vee \exists h \exists \sigma''(\sigma, f)P_M(K')(\sigma'', h) \wedge f = f' \wedge g = g'$
- $(\sigma, f, g)P_M(\pi : x, \sharp K)(\sigma, f', k)$ *iff* $f = f' \wedge$ $\langle f(\pi) f(x), V_M(\sharp K, f, g, \sigma)\rangle \in Say_{w_\sigma, M} \wedge$ $\forall \sigma' \in V_M(\sharp K, f, g, \sigma)(f_w \subseteq g'_w \wedge (\sigma', f \cup g, f \cup g)P_K(\sigma', k, k)$
- $V_M(\sharp K, f, g, \sigma) = \{\sigma' : \ \exists k(\sigma', f \cup g, f \cup g)P_M(K)(\sigma', k, k))\}$

In definition 5, we make use of the glue logic G by means of which we compute the discourse relation used to attach the new information to an available attachment point. This logic exploits the information transferfunction μ from speech acts to information about those speech acts (expressed in the language L_G). μ is not part of G but rather part of our conception of discourse context.

The question as to whether SDRT$_3$ engenders r.e. transitions between states is an important but delicate one. Because we have complicated the notion of contexts, we no longer have a simple correspondence between a first order or higher order formula and a state transition labelled by a DRS (or SDRS). But our contexts can be broken into an intensional part and a part governed by a formula that captures the relation between the information packaging of the two contexts. The language of the relation governing the information packaging contains names of formulas and functions from formulas to the set of free variables in it, together with the function μ that takes us from a formula to lexical information associated with its non-logical constants. Suppose that the function μ gives information from lexical look-up; then μ is clearly computable.

Given that G is decidable and that μ involves only lexical look-up, the relational expression between the information packaging of the input and output contexts (viz. the relation between input set X of speech act discourse referents, current discourse referent and strict partial order on X and the corresponding output elements) is also *decidable.*

The information packaging language L_1 is L_G together with the relation symbol T. L_1 contains names of L formulas as well as names of variables, and variables for sets of variables. It also contains a symbol representing the strict partial ordering on sets of variables. T encodes the transitions concerning the structure of SDRSs as they are modified by new information. To turn the definition of L_G satisfaction into a definition of L_1 *satisfaction*, we need only add a clause for T. Given that every SDRS has a translation into a formula of higher order intensional logic, we can also use T to characterize the structure of such a formula. Suppose that T encodes that transition between structures relevant to determining the sort of information packaging we have supposed in the contexts defined in SDRT$_3$. Calculating this transition involves G, but since G is decidable and determining the extension of T just involves the manipulation of finite structures, L_1 validity is also decidable.

We now translate an SDRS$_3$ context into a pair of formulas—one a formula of intensional higher order logic as before and the other a formula of L_1 that encodes the structure of that context—including of course a set of speech act discourse referents X.

Definition
(χ, δ) is a *characterizing pair of formulas* iff χ is a formula of higher order intensional logic and δ is a formula of L_1 such that $\forall x \in X_\delta \mu(x)$ is L_1 satisfied by the discourse context $\mathcal{P}(K_\chi)[\sigma_0^3]$, where K_χ is the SDRS that is the translation of χ.

Characterization Lemma for SDRT_3:

- For every pair of characterizing formulas (χ, δ) such that χ has free variables U, there is an SDRS (U, C) such that $\mathcal{P}(U, C)[\sigma_0^3] = \{(M, f) : Dom(f) = U \textit{ and } M \models \chi[f]\}$ and $(U, C)L_1$ satisfies δ.
- Every $\text{SDRS}_3(U, C)$ has a pair of characterizing formulas (χ, T) such that: χ has a set of free variables U and $\mathcal{P}(U, C)[\sigma_0^3] = \{(M, f) : Dom(f) = U \textit{ and } M \models \chi[f]\}$ and δ is L_1 satisfied by (U, C)

We use the characterization lemma to define state transitions labelled by SDRSs by keeping the two parts apart. If we can keep L_1 validity decidable (which means keeping G decidable), then we can show that the transitions of SDRT_3 also remain r.e. for the same reasons as in fact 3. But how do these state transitions compare to our previously defined state transitions? Because of the information packaging formula, the transitions are much more restrictively defined than in $\text{SDRT}_{0,1}$; e.g., two logically equivalent SDRSs may not induce the same state transition in SDRT_3.

This sensitivity of the transitions to packaging has advantages I have already mentioned. There are natural language constructions like attitude reports whose semantics is plausibly sensitive not only to truth conditional content but to information packaging as well—thus blurring the distinction with which we set out in the paper. This leads to the thought that the *entire* structure of the SDRS is relevant to information packaging and even occasionally to content. And this suggests one other view of discourse structure—the one originally proposed in Asher (1993).

So far we have been interested in constructing the minimal notions of context needed to attack certain problems and phenomena associated with a theory of the pragmatic-semantic interface. I haven't said anything about the interpretation of certain intensional predicates—predicates of propositions. Candidates for trouble are the predicates for truth, belief and other propositional attitudes. The paradoxes can easily be reintroduced within this framework by predicates that exploit both information packaging and model theoretic content in their semantics. Self-referential paradoxes as of now do not arise, because we have split the information packaging off from the model-theoretic content. The original SDRT of Asher (1993) grew out of a concern with a semantics for abstract objects of the sort needed for an adequate theory of belief reports. This interest in propositions yields a collection of objects in which information packaging and information content are intertwined. Consequently, that theory did not make make the split between information packaging and contents (at least not explicitly) but took the values of $\sharp K$ to be SDRSs themselves. We could call that theory $SDRT_4$. Models for $SDRT_4$ involved techniques familiar to those analyzing the paradoxes—e.g., Frege structure. $SDRT_4$ can also be understood as defining a CCP, but unlike that of our earlier theories it can be shown to be nonaxiomatizable (Asher 1993).

9 Conclusion

In the approach to a variety of systems for discourse interpretation vand problems in discourse interpretation that I have sketched here, I understand the distinction between information packaging and information content to be an essential one. Both content and packaging are needed to understand discourse interpretation, but they should be kept separate if we hope to have a computationally tractable approach to the construction of discourse representations. The distinction between information packaging and content is, however, also a logical distinction that gives us, as the last paragraphs suggest, a slightly different way of thinking about the self-referential paradoxes and about certain problematic constructions in semantics. But both packaging and content are parts of discourse meaning, and this strongly suggests that there are more levels to discourse meaning than just syntax and just model-theoretic semantics.

10 References

Peter Aczel [1988] *Non-well-founded sets.* Center for the Study of Lnaguage and Information. Lecture Notes, no. 14, Stanford California.

Nicholas Asher [1986] Belief in Discourse Representation Theory, *Journal of Philosophical Logic*, **15**, pp.127–189.

Nicholas Asher [1987] A Typology for Attitude Verbs, *Linguistics and Philosophy*, **10**, pp.125–197.

Nicholas Asher [1993] *Reference to Abstract Objects in Discourse*, Kluwer Academic Publishers.

Nicholas Asher [1995] Commonsense Entailment: A Logic for Some Conditionals, in G. Crocco, L. Farinas del Cerro, A. Herzig (eds.), *Conditionals* *and Artificial Intelligence*, Oxford: Oxford University Press, 1995, pp.103-147.

Nicholas Asher [1996] The Mathematical Foundations of Discourse Interpretation, in P. Dekker, M. Stokhof (eds.), *Proceedings of the Tenth Amsterdam Colloquium in Formal semantics*, ITLI lecture notes series, University of Amsterdam.

Nicholas Asher, Michel Aurnague, Myriam Bras, Pierre Sablayrolles, and Laure Vieu [1995] De l'espace-temps dans l'analyse du discours, *Semiotiques* vol. 9, pp. 11-62.

Nicholas Asher and Tim Fernando [1997] Effective Labelling for Disambiguation, *Proceedings of the Second International Workshop in Computational Linguistics*, Tilburg, the Netherlands.

Nicholas Asher and Hans Kamp [1989] Self-Reference, Attitudes and Paradox: Type-Free Logic and Semantics for the Attitudes, in G. Chierchia, B. Partee and R. Turner, (eds.) *Properties Types and Meanings*, Kluwer Academic Publishers, 1989, pp. 85-159.

Nicholas Asher and Alex Lascarides [1994] Intentions and Information in Discourse, in *Proceedings of the 32nd Annual Meeting of the Association of Computational Linguistics (ACL94)*, pp.35–41, Las Cruces.

Nicholas Asher and Alex Lascarides [1995] Lexical Disambiguation in a Discourse Context, *Journal of Semantics*, **12**, pp.69–108.

Nicholas Asher and Alex Lascarides [in press] Bridging, *Journal of Semantics*,

Nicholas Asher and Michael Morreau [1991] Commonsense Entailment: A Modal, Nonmonotonic Theory of Reasoning, *Proceedings of IJCAI 91*, San Mateo, CA: Morgan Kaufman Press, pp.387-392.

Jon Barwise and John Perry [1983] *Situations and Attitudes*, Cambridge, MA: MIT Press.

David Beaver [1994], An Infinite Number of Monkeys, Technical Report, ILLC, Universiteit van Amsterdam.

David Beaver [1996], Presupposition, in J. van Bentham A. ter Meulen (eds.), *Handbook of Logic and Linguistics*, Amsterdam: Elsevier Publishers, pp. 939-1008.

Johann Bos, J., A-M. Mineur and P. Buitelaar [1995] Bridging as coercive accommodation, technical report number 52, Department of Computational Linguistics, Universität Saarbruücken.

Myriam Bras and Nicholas Asher [1994] Le raisonnement non monotone dans la construction de la structure temporelle de textes en franais," *Reconnaissance des Formes et Intelligence Artificielle* (RFIA), Paris, 1994.

Nino Cocchiarella [1989] Philosohpical Perspectives on Formal Theories of Predication, *Handbook of Philosophical Logic*, eds. D. Gabbay and F. Guenthner, Dordrecht: Reidel, pp.253-326.

Mark Danburg-Wyld [forthcoming] *Speech Act Theory and Discourse Representation Theory*, Ph.D. thesis, University of Texas at Austin.

David Dowty [1979] *Word Meaning and Montague Grammar*, Dordrecht: Reidel.

Tim Fernando [1994] What is a DRS?, *Proceedings of the First International Workshop on Computational Semantics*, Tilburg.

Anette Frank [1997] *Context Dependence in Modal Constructions*, Ph.D. Thesis, University of Stuttgart.

Daniel Gallin [1975]: *Intensional and Higher-Order Modal Logic*, Amsterdam: North Holland Publishing Co.

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

Jerry Hobbs, Mark Stickel, Doug Appelt, Paul Martin [1993] Interpretation as Abduction, *Artificial Intelligence*, **63**, pp.69–142.

Hans Kamp [1981] A Theory of Truth and Semantic Representation, in J. Groenendijk, T. Janssen and M. Stokhof (eds.), *Formal Methods in the Study of Language*, Mathematics Center Tracts, Amsterdam.

Hans Kamp [1990] Propositional Attitudes, in C. Anthony Anderson, Joseph Owens (eds.), *The Role of Content in Logic, Language, and Mind*, CSLI Lecture Notes **20**, University of Chicago Press, pp.27–90.

Hans Kamp and Uwe Reyle [1993] *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, Kluwer Academic Publishers.

David Kaplan [1977] *Demonstratives*, John Locke Lectures, Oxford: Oxford University Press.

Alex Lascarides and Nicholas Asher [1993a] Temporal Interpretation, Discourse Relations and Commonsense Entailment, in *Linguistics and Philosophy*, **16**, pp.437–493.

Alex Lascarides and Nicholas Asher [1993b] Semantics and Pragmatics for the Pluperfect, *Proceedings of the European Chapter of the Association for Computational Linguistics* (EACL 93), pp. 250-259, Utrecht, the Netherlands.

William C Mann and Sandra Thompson [1986] Rhetorical Structure Theory: Description and Construction of Text, ISI technical report number RS-86-174.

Richard Montague [1974] *Formal Philosophy*, New Haven: Yale University Press.

Michael Morreau [1995] Allowed Arguments *Proceedings of IJCAI 95*, San Mateo, CA: Morgan Kaufman Press, pp. 1466-1472.

Reinhard Muskens [1996] Combining Montague semantics and discourse representation. *Linguistics and Philosophy*, 19(2).

David Park [1988] Conurrency and automata on infinite sequences, in P. Deussen (ed.), *Proceedings of the 5th GI Conference*, Lecture Notes in Computer Science, vol. 104, Berlin: Springer Verlag, pp. 167–183.

John Perry and Mark Crimmins [1993] *Journal of Philosophy.*

Barbara Partee [1973] Some Structural Analogies Between Tenses and Pronouns in English, *Journal of Philosophy* **70**, pp. 601-609.

Livia Polanyi [1985] A Theory of Discourse Structure and Discourse Coherence. In *Papers from the General Session at the Twenty-First Regional Meeting of the Chicago Linguistics Society*, pp. 25–27.

Robert van Rooy [1997] *Attitudes and Changing Contexts*, Ph.D. Thesis, University of Stuttgart.

Rob van der Sandt [1992] Presupposition Projection as Anaphora Resolution, *Journal of Semantics*, **19**(4).

Robert Stalnaker, R. [1978] Assertion, in Cole, Peter (ed.) *Syntax and Semantics volume 9: Pragmatics*, Academic Press.

Vallduví, E. [1990] *The Informational Component*, PhD. thesis, University of Pennsylvania, Philadelphia.

Author address

Department of Philosophy
The University of Texas
Austin TX 78712-188
USA
email: nasher@mail.la.utexas.edu