Annotating Discourse Structures for Robust Semantic Interpretation

Jason Baldridge Alex Lascarides

1 Introduction

Work in data-intensive linguistics has moved toward machine learning deeper and richer representations of natural language phenomena. Early syntactic annotations were based on simple context-free grammar analyses. Now, richer syntactic models which support the construction of semantic representations are also being used. For example, the Redwoods treebank (Oepen *et al.*, 2002) contains syntactic and semantic analyses from a broad-coverage, hand-crafted Head-driven Phrase Structure Grammar for English sentences from the appointment scheduling and travel planning domains of Verbmobil.

For Redwoods and other similar treebanks, only the compositional semantics of individual sentences is retrievable. Of course, the interpretation of an utterance is influenced by discourse context and extra-linguistic information, in addition to syntax. Our overall goal is to construct a robust statistical model of discourse interpretation that (a) integrates fully with the representations of the compositional semantics of the sentences; and (b) augments them with pragmatic information such as anaphora resolution and communicative goals.

Work has already been done on learning discourse-level semantic information such as co-reference (e.g., Kennedy and Boguraev (1996)) and bridging relations for definite descriptions (e.g. Vieira and Poesio (2000)). Annotated material and algorithms for discourse segmentation and parsing have also been created using a variety of frameworks, such as the Unified Linguistic Discourse Model (Polanyi *et al.*, 2004), Discourse Lexicalized Tree-adjoining Grammar (Webber and Joshi, 1998), and Rhetorical Structure Theory (RST) (Carlson *et al.*, 2001). These frameworks provide varying levels of detail and differing perspectives on how discourse structure should be represented and how discourse segments are related. However, none of them provides a precise semantics for how discourse-level information affects the contextual interpretation of individual utterances.

Other work has focused on plan recognition (Grosz and Sidner, 1990; Litman and Allen, 1990; Green and Carberry, 1994; Lochbaum, 1998), for which the interpretive effect of discourse on an utterance is computed by reasoning about both its logical form and communicative goals. This approach requires complex reasoning about the beliefs, desires and intentions of dialog participants. Additionally, obligations have formed the basis for implementing dialog managers

(Poesio and Traum, 1997). This approach assumes an information state update model of discourse interpretation. It provides deep representations, but is not modular: the information states contain all information from the compositional semantics to speech acts, and no explicit discourse structures are encoded.

Segmented Discourse Representation Theory (SDRT, Asher and Lascarides (2003)) is another approach based on information states. With SDRT, the interpretation of a dialog is influenced by information from lexical semantics, compositional semantics, discourse structure, and the beliefs and goals of dialog participants. SDRT integrates these various information sources in a precise and modular fashion. As such, it avoids some of the complexity associated with reasoning about beliefs and goals by, for example, relying on information sources that are more directly available whenever possible. A crucial part of this is SDRT's rich inventory of rhetorical relations that hierarchically connect discourse segments. But the resulting structures are not trivial to construct, and this has impeded the creation of robust and efficient systems based on SDRT.

All these approaches enable richer interpretations from discourse. However, no one has yet demonstrated the feasibility of acquiring a *comprehensive* model of discourse interpretation that integrates seamlessly with representations of compositional semantics for individual sentences. Current approaches are either too imprecise in terms of the semantic import they assign to discourse-level information, or too complex to represent and compute efficiently.

Our approach to practical discourse interpretation begins with a reduced form of SDRT. Because SDRT analyses are created by referencing multiple, possibly redundant information sources, a large part of an analysis is encoded via discourse segmentation and rhetorical relations alone. SDRT's semantic interpretation of this yields information about the contextually-determined content of utterances. Resolving discourse structure is thus an important step in producing the full contextually-determined information states for dialogs. However, producing full-fledged SDRT discourse structures is complex. To simplify this task, we define a tree-form of discourse structures which can be created more efficiently, both by human annotators and discourse parsers. As such, our annotations are part of a processing strategy – they are not yet usable for linguistic investigations, such as studying the distribution of discourse relations.

2 The Task

The domain under consideration is English appointment scheduling dialogs of Verbmobil. These dialogs form part of the Redwoods treebank, which provides syntactic and semantic analyses for nearly all the utterances. The ultimate goal is to construct a system for building (partial) semantic models of the dialogs. Since the goal of the dialog participants is to agree on a meeting time, the focus for these models is on resolving the denotation of temporal terms and identifying the communicative goals of the participants at each stage in the discourse, including the final agreed upon meeting time.

Consider the following dialog extract from the Verbmobil corpus:[1]

(1) π_{149} PAM: *maybe we can get together, and, discuss, the planning, for say, two hours, in the next, couple weeks,*
π_{150} PAM: *let me know what your schedule is like.*
π_{151} CAE: *okay, let me see.*
π_{152} CAE: *twenty,*
π_{153} CAE: *actually, July twenty sixth and twenty seventh looks good,*
π_{154} CAE: *the twenty sixth afternoon,*
π_{155} CAE: *or the twenty seventh, before three p.m., geez.*
π_{156} CAE: *I am out of town the thirtieth through the,*
π_{157} CAE: *the third, I am in San Francisco.*

This example exhibits several things that a robust model of discourse interpretation should capture. It must resolve the anaphoric temporal description in π_{154} to the twenty sixth of *July* in the afternoon. It must identify both that time and before 3PM on the twenty-seventh as potential times to meet, while ruling out July thirtieth to August third. Additionally, the model should gracefully handle incomplete or ungrammatical utterances like π_{152} and recognize that utterances like π_{151} and π_{152} have no overall effect on finding the time and place to meet.

To achieve both depth and robustness, we apply a mixture of symbolic and statistical methods to create a system for interpreting such dialogs. A hand-crafted grammar is used to analyze utterances, and statistical parse selection mechanisms attempt to identify the best analysis. From these analyses, we obtain interpretations for each utterance. The task is then to update the information in those analyses based on the discourse context. For this we use SDRT. The modularity of SDRT is crucial in this endeavor – our strategy is to create statistical discourse parsers that can robustly provide input for symbolic components, such as theorem provers, that can then construct rich information states. This requires annotating the discourse segmentation and relations of the dialogs in order to train such parsers, a task that is far more practical than *manually* annotating a corpus with its semantic models *directly*—including annotations for the resolutions of all the anaphora and the communicative goals.

We thus intend to obtain semantic models via discourse structures which have a model-theoretical interpretation. SDRT's discourse structures have a precise dynamic semantic interpretation that yields semantic content over and above the compositional semantics of the clauses. The semantic representations of SDRT are fully compatible with representations in the Redwoods treebank, thereby facilitating the integration of semantics from the clause and discourse levels. These properties are exploited in RUDI (Schlangen and Lascarides (2002)), an existing system for computing the semantic consequences of SDRT's discourse structures in determining temporal anaphora and communicative goals for appointment scheduling dialogs. Though RUDI has good coverage on the link from discourse structure to dialog content, it fails to construct discourse structures in a robust manner – the strategy we describe in this paper fills this gap

[1] The sentence numbers are taken from the Redwoods numbering, and the segmentation of the utterances to sentences is also that assumed in Redwoods.

by allowing the creation of robust discourse parsers that produce structures that are directly compatible with RUDI. RUDI can then be used to obtain and annotate the semantic models for the dialogs, thus enabling different strategies for machine learning those models.

3 Segmented Discourse Representation Theory

SDRT provides both a logic for representing (and interpreting) the logical forms of discourse, and a logic for constructing those logical forms. In this paper, we focus on the representations and define a reduced form of them which facilitates annotating them and creating robust methods for constructing them.

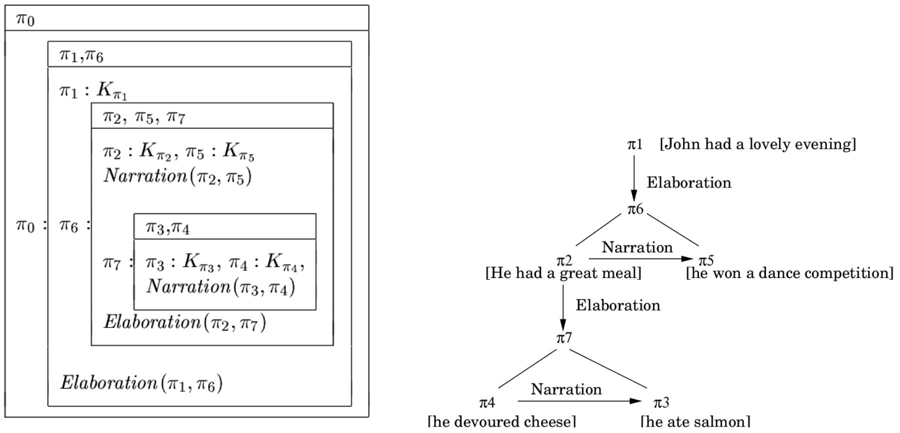
Logical forms in SDRT are interpreted within the framework of dynamic semantics (van Eijk and Kamp, 1997). SDRT extends prior work in dynamic semantics by assuming richer logical forms which feature rhetorical relations such as *Explanation* and *Contrast*. Asher and Lascarides (2003) argue that the rhetorical structure of a discourse has an impact on its semantics—particularly the interpretation of anaphora—that cannot be modeled straightforwardly through other information sources such as world knowledge or reasoning about beliefs and intentions. Prior work, such as Rhetorical Structure Theory (RST, Mann and Thompson (1988)), has developed sets of rhetorical relations that hold between utterances; SDRT extends this by supplying rhetorical relations with a precise dynamic semantics. This dynamic semantic interpretation explains how the content of a discourse augments the compositional semantics of its clauses.

SDRT adds two new expressions to standard logical forms in dynamic semantics (e.g., DRSs from DRT, Kamp and Reyle (1993)). *Speech act discourse referents* label content (either of a clause or of text segments) to keep track of the token utterances in the discourse; and *rhetorical relations* such as *Continuation* and *Explanation* relate these referents. The resulting structures are referred to as segmented DRSs or SDRSs. An SDRS consists of two things: a set of speech act discourse referents (usually written $\pi_1, \pi_2, \ldots$) and a mapping from those referents to SDRS-formal. These formulas are comprised of the well-formed formulas for representing the logical forms for clauses, as well as formulas such as $Explanation(\pi_1, \pi_2)$ involving rhetorical relations.

As an example, consider (2) and its DRT-style representation in Figure 1(a).

(2) π_1 *John had a great evening last night.*
π_2 *He had a great meal.*
π_3 *He ate salmon.*
π_4 *He devoured lots of cheese.*
π_5 *He then won a dancing competition.*

In Figure 1(a), the semantics of each discourse unit π_i is abbreviated as K_{π_i}. The discourse relations then connect these units and lead to further content. For example, that the event described by π_2 temporally preceded that of π_5 follows from the model-theoretic interpretation of $Narration(\pi_2, \pi_5)$. Other relations



(a) (b)

Figure 1: Discourse (2) represented (a) in DRT-style notation and (b) as a graph.

lead to the reverse temporal precedence; for example, *Explanation*(π_1, π_2) indicates that the event in π_1 cannot have preceded that of π_2. It also follows from *Explanation*(π_1, π_2) that the event in π_2 caused that of π_1.

There are other ways in which this information can be represented, so as to show the discourse structure more clearly. Figure 1(b) conveys the same information as Figure 1(a), but it also shows the rhetorical relations and segmentation as well as which relations are coordinating and which are subordinating: the former are indicated by horizontal lines and the latter by vertical ones. This distinction is important since it affects constraints on anaphora. For example, the coordinating relation *Narration*(π_2, π_5) blocks anaphoric reference to the cheese in π_4, so it is not possible to follow π_5 by *It was Wensleydale* with the intended referent of *It* being the cheese which John devoured.

4 Non-treeness in discourse structure

SDRT does not restrict discourse structures to trees, contrary to what is typically assumed in other approaches, such as RST. Thus, a given utterance (or segment) can take part in multiple relations; or equivalently, can be used to perform more than one illocutionary act (Moore and Pollack, 1992). Asher and Lascarides (2003) show that this is necessary in a competence theory of discourse understanding. However, this property does nonetheless complicate the creation of efficient discourse parsers. Here, we describe four (mutually compatible) ways in which non-treeness can arise, and in the following section, we provide a tree-approximation of SDRSs that can be efficiently processed without losing all of the non-tree aspects of full SDRSs.

The first manner in which non-treeness arises is when two utterances stand in multiple relations to each other. For example, *B*'s utterance in (3) is both an indirect answer to A's question and an elaboration of a plan to achieve the goal behind A's question. Thus both $IQAP(\pi_1, \pi_2)$[2] and *Plan-Elab*(π_1, π_2) hold.

(3) π_1 A: *Can we meet next week?*
π_2 B: *I'm free on Tuesday.*

This contrasts with a response to π_1 by B such as *I'll be on vacation.* This is still an indirect answer, but the plan behind A's question is rejected, rather than adopted and elaborated on. This illocutionary move is conveyed by the relation *Plan-Correction*, and thus *Plan-Correction*(π_1, π_2) and $IQAP(\pi_1, \pi_2)$ hold.

Another way in which discourse structures depart from trees is that multiple utterances attach, via different relations, to a single utterance. For example, in (2), π_2 is related by *Elaboration* to π_3 and by *Narration* to π_5.

Conversely, an utterance can attach to multiple utterances in the context. This indicates that it makes more than one illocutionary contribution to more than one part of the discourse context. For example, an adequate interpretation of dialog (4) requires *both Counterevidence*(π_2, π_3) and *Elaboration*(π_1, π_3) to be identified.

(4) π_1 A: *Max owns several classic cars.*
π_2 B: *No he doesn't.*
π_3 A: *He owns two 1967 Alfa spiders.*

Logically inferring one of these relations is co-dependent on inferring the other. *Elaboration*(π_1, π_3) semantically entails that 1967 Alfa spiders are classic cars, which is a necessary condition for *Counterevidence*(π_2, π_3) holding.

The above three departures from treeness are all local in nature. A distinction between the content level and the intentional level (Moore and Pollack, 1992) can cause more global divergence. For example, consider the following:

(5) π_{385} KCF: *How about, Wednesday, the fourth.*
π_{386} KCF: *In the afternoon?*
π_{387} KCF: *Maybe around two?*
π_{388} JCB: *Wednesday is okay,*

At the content level, the speaker KCF follows π_{385} with two questions elaborating the specification of a meeting time. At the intentional level, these questions elaborate a plan to reach a meeting time. In SDRT terms, the former is represented with the relation *Elaboration*$_{qq}$ and the latter with *Q-Elab*. The semantics of *Elaboration*$_{qq}(\pi_{385}, \pi_{386})$ specifies that the content of π_{386} is a sub-part of that of π_{385}. The semantics of *Q-Elab*(π_{385}, π_{386}) instead specifies that any answer to π_{386} augments a plan to achieve the communicative goal behind π_{385}.

In this case and other such examples in our corpus, these two alternatives encode a spurious ambiguity. This is a crucial factor in allowing the strategy we describe in the next section to work for our domain. The situations described by Moore and Pollack (1992) are not spurious. There, both the content and intentional levels must be represented to encode the full import of the discourse.

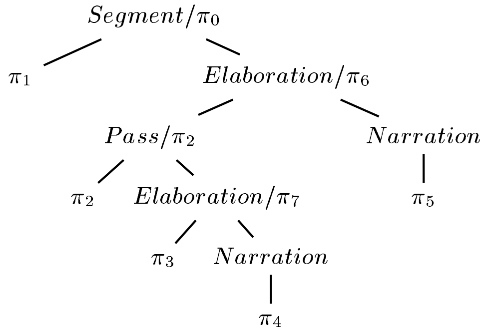
[2] $IQAP$ stands for *indirect question answer pair.*

5 Tree Approximation of SDRSs

The previous section outlines cases where trees are insufficient to fully represent discourse structure – graphs seem necessary for a competence theory of discourse understanding. However, such graphs are computationally expensive to compute and complex to annotate. Here, we define a tree-based approximation that encodes nearly all of the information in SDRSs and can be created more efficiently.

The graph in Figure 1(b) uses utterances and segments as node labels. Graph arcs are labeled with the relations which connect them. This is an intuitive representation, but other strategies can be used. We can instead represent SDRSs with trees that establishes dependencies between utterances, as with syntax trees. For example, consider the following tree SDRS for dialog (2).

(6)



In this tree, utterances are leaves, and the relations indicate different kinds of phrase labels for discourse segments. Each node in (6) establishes dependencies between a *head* segment and other segments which are descendants of the node. This is done via discourse referents associated with each node. A leaf node's discourse referent is that of the utterance it represents. If a non-terminal has a segment label (e.g., the *Elaboration* node with π_6 in (6)), that is the discourse referent; otherwise, the referent is that of the node's head daughter. The non-rhetorical relation *Segment* creates left-branching subsegments and creates a new referent. The other non-relation *Pass* is similar, except that it passes up the referent of its head segment so that it can take part in relations higher up in the tree. Each node must have one and only one leaf, *Segment*, or *Pass* daughter; this daughter is the node's head.

Relations are recovered by inspecting a node and taking the label as the relation, the discourse referent of its parent as the first argument, and its own discourse referent as the second argument. For example, we get *Narration*(π_3, π_4) as one of the relations represented in (6). The *Pass* nodes allows us to naturally represent the graph-inducing situations where multiple utterances attach to a single utterance. For example, π_2 in (2) stands in an *Elaboration* relation with the segment π_7 containing π_3 and π_4, and a *Narration* relation with π_5.

The tree in (6) represents the following set of relations: *Elaboration*(π_1, π_6), *Elaboration*(π_2, π_7), *Narration*(π_3, π_4), *Narration*(π_2, π_5). The scope of the labels –as indicated by the boxes in Figure 1(a)– must also be recovered. This is easily done: for any node which introduces a segment label, we create an *outscopes* relation between that label and the discourse referents of the node's daughters. This gives us $outscopes(\pi_0, \pi_1)$, $outscopes(\pi_0, \pi_6)$, $outscopes(\pi_6, \pi_2)$, $outscopes(\pi_6, \pi_7)$, $outscopes(\pi_6, \pi_5)$, $outscopes(\pi_7, \pi_3)$, $outscopes(\pi_7, \pi_4)$. Together, these encode the same information as the representations in Figure 1. The tree in (6) thus completely determines the SDRS for discourse (2).

Consider again example (3), in which two utterances are connected by two relations, *Plan-Elab*(π_1, π_2) and $IQAP(\pi_1, \pi_2)$. This is typically represented as two arcs between π_1 and π_2 in a graph, but the situation can also be handled by a single combined relation *Plan-Elab*/$IQAP(\pi_1, \pi_2)$. In our tree approximation, this is represented with a node *Plan-Elab/IQAP* that is expanded into the two separate relations when the SDRS logical form is produced from the tree.

We can thus encode two types of graph-like properties straightforwardly with trees. It might seem from this that whether we attribute the property of treehood to discourse structures or not depends entirely on the graphical representation that we assume. But this is not the case: the mapping of the SDRSs to this proposed tree-notation will not always fully preserve the discourse structural information in the way (6) does for (2).

Example (4) shows a situation for which we cannot represent all the relations directly with our tree approximations. There, the last utterance π_3 connects to both of the previous utterances. While it is not out of the question that this could be encoded in some manner in a tree (e.g., by allowing utterances to corresponds to multiple leaves), it would almost certainly severely complicate segmentation and the interpretation function that produces SDRSs from them. The reason for this is that the tree form we have designed for discourse centers around the left arguments of rhetorical relations. Heads are always on the left, so *Pass* nodes can only pass up the referent for left segment; such referents can then take part in multiple relations. The topology of the trees could be modified to reverse this situation, but we would then be unable to capture the more common case of multiple uses as the left argument.

This sort of situation does arise in the Verbmobil dialogs, but being unable to identify one of the relations does not typically block the correct model for the dialog from being constructed. For example, consider the following:

(7) π_1 A: *Shall we meet on Wednesday?*
π_2 A: *How about one pm?*
π_3 B: *Would one thirty be OK with you?*

A full SDRS would specify *Plan-Correction*(π_2, π_3), reflecting the fact that meeting at one-thirty and meeting at one are incompatible goals, and *Q-Elab*(π_1, π_3), reflecting the fact that any answer to this question elaborates a plan to meet on Wednesday. Here, the consequences of dropping one of these relations is benign with respect to the interpretation of temporal anaphora and the goals. Together with other information, such as conventional constraints on temporal bridging relations (Schlangen and Lascarides, 2002), the interpretation where *one-thirty* in π_3 resolves to one-thirty on Wednesday and the goal is to meet at one-thirty on Wednesday is derivable from either of these rhetorical attachments.

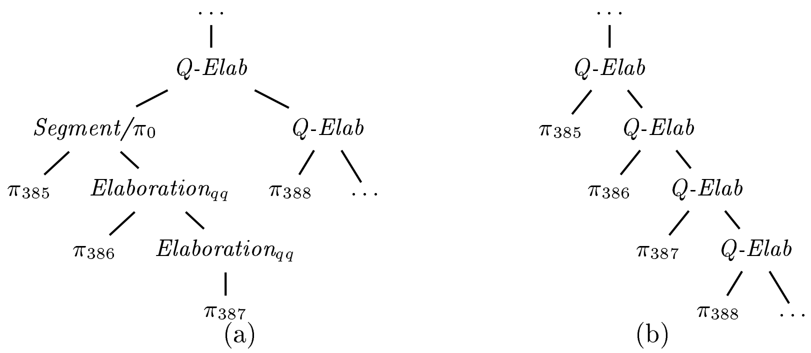


Figure 2: Structure of (5) at the (a) content level and (b) intentional level.

Another thing that we do not attempt to encode in these tree approximations is when the discourse structure at the content level diverges from the intentional level. This was exemplified in (5), for which we have the structure in Figure 2(a) at the content level and in Figure 2(b) at the intentional level. The difference between SDRT's dynamic semantic interpretations of these two structures is that in the former there is a segment π_0 which is absent from the latter. π_0 labels an SDRS-formula whose interpretation can be paraphrased as the single question: *How about we meet on Wednesday afternoon at maybe around two?*. However, assuming that the implementation of SDRT's model theory has complete coverage in terms of the goals that are associated with the rhetorical relations in this scheduling domain, both structures generate the same goal for π_{388} and the same resolution of its temporal anaphora.

We cannot represent both of these structures with a single tree, so a choice must be made when annotating such situations. Our convention is to use the content-level structure in such cases of ambiguity where the speaker does not change, and the intention-level structure when the speaker changes. This maximizes the correspondence between segments in the discourse structure on the one hand and turns in the dialog on the other. It also increases the variety of relations and structures in the training corpus, thereby reducing the risk of sparse data when it comes to training discourse parsers.

Another factor which must be dealt with are utterances which are ignorable for discourse interpretation. These are inserted as leaves of three types: *Irrelevant*, *Pause*, and *Pleasantry*. For example, (1) contains the *Pause* π_{151} and the *Irrelevant* disfluency π_{152}. When producing the SDRS from a tree, these leaves are ignored. However, they cannot just be dropped entirely from the trees themselves — discourse parsing models must have access to examples of such utterances in order to detect and handle them on unseen data. Breaking ignorable utterances into three types is motivated by the expectation that pauses and pleasantries will have some utility in determining segmentation.

6 Annotation

The task of annotating the Verbmobil dialogs is now underway. Discourse structure is encoded with our tree approximations using 35 distinct rhetorical relations plus the non-rhetorical *Segment* and *Pass* relations. Leaves are labeled either with one of three sentence moods –indicative, interrogative, or imperative– or one of three ignorable types –irrelevant, pause, or pleasantry. We assume, for simplicity, the segmentation given by the treebank for determining the elementary discourse units. Usually this assumption is appropriate; when it is not, we note the presence of sub-segments and possible relations between them. These sub-segments do not appear individually in the annotated discourse structure.

In the first stage, 25 randomly selected dialogs comprising 741 utterances were independently annotated by two annotators, after which difficult issues were resolved and conventions were chosen together. In the second stage, ten more randomly selected dialogs comprising 363 utterances were independently annotated to measure annotation time and annotator agreement. For these dialogs, both annotators averaged about 45 seconds per utterance. We feel this is a good rate for the initial annotation for discourse structures, especially given that they are currently input directly as XML without any special purpose annotation tools.

To evaluate inter-annotator agreement, we extract the relations from each discourse tree and measure the overlap as an f-score.[3] Table 8 gives the labeled and unlabeled scores when comparing the two annotators against each other and each annotator against a baseline that assigns completely right-branching structures using the most common relation (*Continue*) as the label for all nodes.

(8)

	Inter-annotator	Baseline - Ann. 1	Baseline - Ann. 2
Labeled	50.2%	12.4%	11.6%
Unlabeled	66.0%	45.5%	42.1%

These figures show that the annotators agree with each other far more reliably than a simple baseline does with either of them. Nonetheless, the inter-annotator agreement would ideally be higher. One reason for the low agreement is the need to establish more conventions regarding spurious ambiguities in discourse structures. In many cases where annotations differ, it would still be possible to recover the correct semantic models from either. Unsurprisingly, another reason is the harshness of the evaluation, which marks two constituents as entirely different even when just a small part diverges. An indication of this can be seen by comparing one-sided f-scores: inter-annotator agreement for matching both the relation and the first argument is 60.0%, and matching both the relation and the second argument is is 63.7%, compared to 50.2% matching all three.

[3] F-score is calculated as $\frac{2 \times precision \times recall}{precision + recall}$. Because our task involves structure building in addition to classification, we do not use the Kappa statistic.

7 Conclusion

The tree representations of SDRSs express quite naturally both rhetorical connections and discourse segmentation, and they do so without the need to manually keep track of semantic indexes and segment labels. The use of trees captures some situations where discourse structure appears to require graphs. Cases where multiple relations hold between two utterances or multiple utterances attach to a single segment are fully recoverable when SDRSs are built from the trees. Nonetheless, the discourse trees are indeed approximations, as demonstrated by the fact that they cannot capture situations where a single utterance attaches to multiple utterances nor the situation where the content and intentional structures diverge. In these cases the full SDRS is not recoverable and hence the complete model-theoretic interpretation may not be either. But for the Verbmobil dialogs this is not problematic in that resolving all the temporal anaphora and computing the time and place at which the participants aim to meet is possible from the dynamic semantic consequences of the partial SDRSs that are recoverable from the tree approximations.

An advantage of using trees approximations is that we can straightforwardly apply probabilistic parsing methodologies for context-free grammars to create and disambiguate discourse structures. We are currently building a discourse parser which estimates its parameters from the annotated dialogs. In the next phase of annotation, we will use the parser to semi-automate the annotation of discourse structures. The SDRSs obtained from the parser will ultimately be used for discourse-situated parse selection and as input to the implementation of SDRT's model theory in RUDI (Schlangen and Lascarides, 2002).

The explicit purpose of our annotations is for training such probabilistic discourse parsers. The fact that the trees do not represent full-fledged SDRSs means that the annotated material cannot yet be used for other purposes such as studying the distributional properties of SDRT relations in the corpus. Nonetheless, since the trees represent a subset of the full SDRSs, the annotations can be directly extended for such purposes later.

Acknowledgments

We would like to thank Ivana Kruijff-Korbayová, Geert-Jan Kruijff, David Schlangen, and the anonymous reviewers. This work was supported by Edinburgh-Stanford Link R36763, ROSIE project.

References

Asher, Nicolas and Lascarides, Alex, 2003. *Logics of Conversation*. Cambridge University Press.

Carlson, L., Marcu, D., and Okurowski, M., 2001. "Building a Discourse-Tagged Corpus in the Framework of Rhetorical Structure Theory." In *Proceedings of the*

2nd SIGDIAL Workshop on Discourse and Dialogue, Eurospeech 2001. Denmark, September 2001.

Green, N. and Carberry, S., 1994. "A Hybrid Reasoning Model for Indirect Answers." In *Proceedings of the 32nd Annual Meeting of the ACL.* Las Cruces, New Mexico, 58–65.

Groenendijk, J. and Stokhof, M., 1991. "Dynamic Predicate Logic." *Linguistics and Philosophy* 14:39–100.

Grosz, B. and Sidner, C., 1990. "Plans for Discourse." In J. Morgan P. R. Cohen and M. Pollack (eds.), *Intentions in Communication*, MIT Press. 365–388.

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

Kennedy, Christopher and Boguraev, Branimir, 1996. "Anaphora for everyone: Pronominal anaphora resolution without a parser." In *Proceedings of COLING 1996.*

Litman, D. and Allen, J., 1990. "Discourse Processing and Commonsense Plans." In J. Morgan P. R. Cohen and M. Pollack (eds.), *Intentions in Communication*, MIT Press. 365–388.

Lochbaum, K., 1998. "A Collaborative Planning Model of Intentional Structure." *Computational Linguistics* 24(4):525–572.

Mann, William C. and Thompson, Sandra A., 1988. "Rhetorical structure theory: Toward a functional theory of text organization." *Text* 8(3):243–281.

Moore, J. D. and Pollack, M. E., 1992. "A Problem for RST: The Need for Multi-Level Discourse Analysis." *Computational Linguistics* 18(4):537–544.

Oepen, Stephan, Toutanova, Kristina, Shieber, Stuart, Manning, Christopher, Flickinger, Dan, and Brants, Thorsten, 2002. "The LinGO Redwoods Treebank: Motivation and Preliminary Applications." In *Proc. of the 19th International Conference on Computational Linguistics.* Taipei, Taiwan.

Poesio, M. and Traum, D., 1997. "Conversational Actions and Discourse Situations." *Computational Intelligence* 13(3).

Polanyi, Liv, van den Berg, Martin, Culy, Chris, Thione, Gian Lorenzo, and Ahn, David, 2004. "A Rule Based Approach to Discourse Parsing." In *Proceedings of SIGDIAL 2004.* Boston, MA.

Schlangen, D. and Lascarides, A., 2002. "Resolving Fragments using Discourse Information." In *Proceedings of the 6th International Workshop on the Semantics and Pragmatics of Dialogue (Edilog).* Edinburgh.

van Eijk, J. and Kamp, H., 1997. "Representing Discourse in Context." In J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Linguistics*, Elsevier. 179–237.

Vieira, Renata and Poesio, Massimo, 2000. "Processing definite descriptions in corpora." In *Corpus-based and computational approaches to anaphora*, UCL Press.

Webber, B. L. and Joshi, A., 1998. "Anchoring a Lexicalized Tree Adjoining Grammar for Discourse." In *Proceeedings of the ACL/COLING Workshop on Discourse Relations and Discourse Markers*. Montreal.