Issues in Linguistic Segmentation

Janyce M. Wiebe
Department of Computer Science and the Computing Research Laboratory
Box 30001/Dept. CS
New Mexico State University
Las Cruces, NM 88003
wiebe@nmsu.edu

This paper addresses discourse structure from the perspective of understanding.

It would perhaps help us understand the nature of discourse relations if we better understood what units of a text can be related to one another. In one major theory of discourse structure, Rhetorical Structure Theory (Mann & Thompson 1988; hereafter simply *RST*), the smallest possible linguistic units that can participate in a rhetorical relation are called *units*, and "are essentially clauses, except that clausal subjects and complements and restricted relative clauses are considered as parts of their host clause units rather than as separate units" [p. 248]. But both Dale and Meteer (in these proceedings) point out that rhetorical relations can appear **within** clausal units. (Dale's argument will be discussed at the end of this paper.) For example, the relation that is expressed in two clauses in (1.2) is expressed in only one clause in (1.1) (from Meteer, these proceedings):

(1) 1.1 My flicking the switch caused the light to turn on.
1.2 Because I flicked the switch, the light turned on.

Similarly, Hwang & Schubert (1992), in their work on recognizing temporal relations among episodes in discourse, argue for a "fine structure" of discourse, in which temporal relations can be established even among episodes of subordinate clauses.

This paper points out another discourse phenomenon that calls for a "fine structure" of discourse. In passages containing *attitude reports*—reports of agents' beliefs, knowledge, intentions, perceptions, etc.—rhetorical relations can hold such that one or more of the linguistic units involved in a relation is only part of a sentence.[1] In some cases, such a unit may be smaller than the smallest possible unit in RST. Specifically, only the **complement** of an attitude report, rather than the entire sentence, might be involved in some particular relation. (An example is (2) below, which will be discussed shortly.) To make matters more concrete, we will consider short passages in which an attitude report participates in a relation indicated by the cue phrase 'but', where 'but' is being used to connect clauses.

[1] Note that this paper is meant to illustrate some complexities that I believe require attention. Undoubtedly, similar discourse structures occur that do not involve attitude reports.

Consider the following passage:

(2) 2.1 John knew that Mary had never been introduced to Sam.
2.2 But she had been introduced to Derek.

One reading of this passage is that (2.2), as well as (2.1), presents an attitude of John. Imagine that (2) appears in a narrative in which John is forming a plan, and whom Mary has and has not met is somehow important to this plan. Under this reading, (2.2) is an example of a sentence that presents the attitude of some agent X, even though X nor the attitude are mentioned in the sentence. (Detecting such sentences specifically in third-person fictional narrative text was the focus of previous work; see Wiebe 1990.)

Notice that 'But' in (2.2) is being used to connect clauses, and not in addition to mark the beginning of a new discourse segment (as the term *discourse segment* is used in Grosz & Sidner 1986). The question we are asking is what clauses are being connected by 'But' in (2)? Under the reading described above, the following are the clauses participating in the relation:

Mary had never been introduced to Sam. [the **complement** of (2.1)]
But she had been introduced to Derek.

Contrast (2) with (3):

(3) 3.1 John thought that Richard had stabbed him in the back.
3.2 But John was often too suspicious of Richard.

In (3), the entire sentences (3.1) and (3.2) participate in the relation indicated by 'But'.

Deciding which linguistic units are involved in a relation is not sufficient for understanding how they are related, of course. Generally speaking, knowledge about the world and/or what the speaker or writer is trying to accomplish in the discourse (Moore & Pollack 1992) would presumably be involved in arriving at the actual contrasts being made in (2) and (3) (or whatever sorts of relations are being indicated by 'But'). But to hope to arrive at an understanding of such texts, an NLU system must entertain the possibility that one or more of the linguistic units involved in the relation may be only a clausal complement.

Following is another short passage in which 'but' indicates a relation involving an attitude report. For this passage, RST units are sufficiently fine-grained. The passage is of interest to us here because the main clause of the second sentence, (4.3), is not involved in the relation (what are numbered in this passage are RST units):

(4) 4.1 The car was finally coming toward him.
4.2 He finished his diagnostic tests, **4.3** feeling relief.
4.4 But then the car started to turn right.

The relation indicated by 'But' is between his relief at the car coming toward him ((4.1) and (4.3) together) and the car then turning right (4.4). That is, for the purpose of understanding the relation indicated by 'But', (4.1) and (4.3) are grouped together, which are in turn grouped with (4.4). But there are clearly also narrative relations to be recognized in this passage (e.g., (4.2) and (4.4) are in a sequence relation), which involve other groupings of the clauses. Many (such as Moore & Pollack 1992, and Dale, Hughes & McCoy, Meteer, and Moser & Moore in these proceedings) have noted that more than one type of relation can simultaneously hold among elements of the discourse; we see another example of this here. In discourses presenting attitudes, because they present states, events, and objects as well as attitudes toward them, a linguistic unit can be involved in more than one kind of relation, possibly grouped with different units.

I used "grouped" above for lack of a completely appropriate term from the literature. It would be good to talk about these groupings as discourse segments of the linguistic structure of Grosz & Sidner's theory, to distinguish linguistic structure from the non-linguistic basis for that structure. This would be misleading, however, because discourse is structured in Grosz & Sidner's theory on the basis of intentions. A rhetorical relation holding between pieces of a discourse, such as the one indicated by 'But' in (4), does not necessarily make them into a discourse segment (see Grosz & Sidner 1986, p. 188, Moore & Paris 1992, p. 46, and Dale, these proceedings, for discussion). Another possible term is Mann & Thompson's *text span*, which suggests (at least in practice) segments that are either units, or composed only of adjacent units (excluding, for example, a segment composed of (4.1) and (4.3) together with (4.4)).

The examples given above do not illustrate the range of possible discourse structures in which attitude reports participate. I am currently analyzing text segments (from on-line texts) that contain both attitude terms and particular cue phrases to try to identify the various possibilities. My goal is to develop a mechanism that uses syntactic and lexical knowledge to identify the segments involved in the relation. Without access to world and/or intentional knowledge, such a mechanism could produce only likely hypotheses; the idea is to see if information extracted by the non-discourse components of an overall NLU system could be used to constrain processing at deeper levels (Wiebe 1990, Bergler to appear, Passonneau & Litman 1993, and Hirschberg & Litman to appear each address using one or more of syntactic, lexical, orthographic, and intonational information to perform discourse tasks).

In summary, in discourses with attitude reports, (A) linguistic units smaller than RST units may be involved in a discourse relation, and (B) a single linguistic unit may be involved in more than one kind of relation, possibly together with different linguistic units. As mentioned above, others have also noted (A) and (B). Dale makes the interesting argument that, among other things, (A) and (B) suggest that we should banish those rhetorical relations which simply mirror underlying knowledge-base relations. In many cases, he points out, rhetorical relations are simply subject-matter relations—in establishing such relations, all we are really doing is identifying knowledge-base relations between entities mentioned in the discourse. With this in mind, the fact that we find many instances of (A) and (B) is not surprising. Since various syntactic constituents evoke various objects, states, and events, it is not surprising that one can find discourse relations (mirroring knowledge-base relations) that involve various pieces of sentences.

I think that Dale makes some very good points. As I suggested above, one finds structures such as the ones illustrated in this paper because discourses can present attitudes towards things as well as presenting those things themselves. Thus, we find relations among parts of sentences evoking only the objects of attitudes, as well as among those evoking the attitudes themselves. Further, I certainly do not disagree with Dale's suggestion that focusing on those rhetorical relations that are clearly not domain relations would be a way to better understand communicative intentions.

Given a rich knowledge base, however, out of all of the possible knowledge-base relations that can hold among all the things evoked in a discourse, only some are intended to be picked out as the basis for coherence. (If we allow default inference, the number of possible relations is astounding.) Hobbs (1979), among others, argues this. As such, certain groupings of linguistic units, i.e., those evoking the things involved in these relations, are more important than others for establishing coherence. Perhaps the ideal discourse model is one in which the process of arriving at these groupings and associated relations is governed by a process of intention recognition. But investigating local coherence directly—how it manifests itself in various contexts in naturally-occurring texts, and how non-pragmatic information might be exploited to recognize it—could provide important constraints for intention-based models.

References

[1] Bergler, S. From Lexical Semantics to Text Analysis (to appear). In P. Saint-Dizier and E. Viegas (eds.), *Computational Lexical Semantics*, Cambridge University Press.

[2] Grosz, B. J. & Sidner, C. (1986). Attention, Intentions, and the Structure of Discourse. *Computational Linguistics* 12(3): 175-204.

[3] Hirschberg, J. and Litman, D. (to appear). Empirical Studies on the Disambiguation of Cue Phrases. *Computational Linguistics*.

[4] Hobbs, J. R. (1979). Coherence and Coreference. *Cognitive Science* 3(1): 67–90.

[5] Hwang, C. H. & Schubert, L. K. (1992). Tense Trees as the "Fine Structure" of Discourse. In *Proc. of ACL-92*: 232–240.

[6] Mann, W. & Thompson, S. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text* 8(3): 243-281.

[7] Moore, J. & Paris, C. (1992). Planning Text for Advisory Dialogues: Capturing Intentional and Rhetorical Information. *Technical Report 92-22*, (Pittsburgh: University of Pittsburgh Dept. of Computer Science).

[8] Moore, J. & Pollack, M. (1992). A Problem for RST: The need for Multi-Level Discourse Analysis. *Computational Linguistics* 18(4): 537-544.

[9] Passonneau, R. & Litman, D. (1993). Intention-Based Segmentation: Human Reliability and Correlation with Linguistic Cues. In *Proc. of ACL-93*.

[10] Wiebe, J. (1990). Recognizing Subjective Sentences: A Computational Investigation of Narrative Text. *Ph.D. dissertation. Technical Report 90-03* (Buffalo: SUNY Buffalo Dept. of Computer Science).