Reference Manual for the Analysis and Annotation of Rhetorical Structure (Version 1.0)*

Brian Reese, Julie Hunter, Nicholas Asher, Pascal Denis and Jason Baldridge
Departments of Linguistics and Philosophy
University of Texas at Austin

June 14, 2007

Contents

1   Introduction

This document is a reference manual for the segmentation and annotation scheme developed for the DISCOR project (Discourse Structure and Coreference Resolution). The goal of DISCOR is to test certain hypotheses about the interaction between discourse structure and the resolution of anaphoric links in a text. Since Polyani [REFERENCE] many have assumed that discourse structure will constrain the set of accessible antecedents to an anaphoric expression via a "right frontier" contstraint.

The annotation scheme presented in this document was used to annotate the MUC6 and ACE2 corpora for discourse structure, following the model of discourse structure hypothesized by Segmented Discourse Representation Theory (Asher and Lascarides 2003). These corpora were chosen because they have already been annotated for co-reference To date, 60 Wall Street Journal articles from the MUC6 corpus have been annotated, and annotation of the newswire section of the ACE2 corpus has commenced. We hope that this corpus of discourse structure annotated texts will be made available to the larger community for research in other NLP tasks, such as text summarization, etc., and will thereby supplement existing discourse corpora such as the RST corpus (Carlson et al. 2003) and the Penn Discourse Treebank (The PDTB Research Group 2006).

The DISCOR annotation scheme employs a relatively small vocabulary of discourse relations: 14 in total. In contrast, the RST corpus uses 78 relations, although these relations can be grouped to form a more coarse-grained set of 16 classes based on overall semantic similarity. Wolf and Gibson (2005) use 11 relations.

There are two steps in the annotation process: segmentation and annotation. §2 details the conventions for defining the elementary discourse units of a given text, while §3 describes the semantics of the 14 discourse relations used in DISCOR. §4 explains the annotation process and conventions, and also provides background on SDRT.

2 Segmentation

The first step when constructing of a corpus of discourse structure annotated texts is to identify the elementary discourse units (EDUs), or "words", of a given text. Segmentation of EDUs is currently done manually, but an automatic text segmenter is being developed. We turn now to a detailed discussion of our segmentation conventions. '<***>' is used to mark EDU boundaries.

2.1 Base Case

A segment boundary is placed after all punctuation marking the end of a sentence, i.e., periods, question marks, and exclamation points.

2.2 Sentence Internal Segmentation

In general, subsentential constituents are treated as EDUs if they serve a discernible discourse function. There are, however, two exceptional cases. First, if part of a sentence serves a discernible discourse function but segmenting that part results in a discontinuous segment, then we do not segment that part of the sentence. For example, in the following sentence the clause beginning with 'which' supplies background information on Arnold. Normally we segment clauses beginning with 'wh-' words, as we explain below. In this case, however, segmenting the 'wh-' clause would split the larger EDU 'Privately held Arnold...handles advertising for such major corporations as McDonald's Corp.,...' into two smaller units that do not qualify as EDUS according to our rules. Thus we do not segment.

(1)   Privately held Arnold, which had about $750 million
       in billings and $90.7 million in revenue last year,
       handles advertising for such major corporations as
       McDonald's Corp., ...

Second, if part of a sentence serves a discernible discourse function but is embedded under information that cannot be segmented, then we do not segment. The following sentence contains a conditional clause, "without pressure he will not respond" which would normally be segmented as "without pressure¡***¿he will not respond". Yet because this conditional is in the scope of "understand", which we do not segment according to our rules, we do not segment the conditional clause in this case.

(2)   "What we want is to be leading the group to understand
       that without pressure he [Milosevic] will not respond."

Punctuation and discourse markers are good surface syntactic cues for detecting sentence internal EDUs. Segment boundaries are placed *after* punctuation—including periods, commas, hyphens, colons and semi-colons—but *before* discourse connectors such as "and," "or," etc. and complementizers such as "that," "if," "whether," etc.

2.2.1 Subordinate Clauses

Many, but certainly not all, cases of syntactic subordination introduce a new elementary discourse unit. Complements of verbs of communication, for example, introduce EDUs.

- Verbs of communication:
  acknowledge, add, announce, argue, concede, disclose, explain, indicate (with animate agent only), make it clear, note, point out, propose, remark, report, respond, say, tell, testify, etc,

The complements of attitude verbs do not introduce discourse segments unless there is discourse structure within the subordinate clause. In (26), for example, the complement of *predict* is segmented because the 'by' clause explains how Milosovic will respond.

(3) But European allies predict<***> that Milosevic will respond to encouraging gestures<***> by taking further steps toward peace.

2.2.2 Relative Clauses

We assume that nonrestrictive relative clauses introduce EDUs, unless doing so results in a discontinuous EDU (which is often the case). Restrictive relative clauses are never segmented as they do not serve a discourse function but rather restrict the denotation of a common noun. Under these conventions, the following segmentation is correct.

(4) First, we go to Jane Clayson,<***> who has the details of the accident.

The segmentations in (5-a) and (5-b), on the other hand, are incorrect. In (5-a), segmenting the nonrestrictive relative clause results in a discontinuous EDU. In (5-b), the relative clause restricts the denotation of the common noun *fact*.

(5)   a. The president of the Association of Professional
           Flight Attendants,<***> which represents Americans's
           more than 10,000 flight attendants,<***> called
           the request for mediation...
       b. Contrary to the implication in your article, the
           fact<***> that the outside directors have employed
           independent legal and financial counsel in response
           to the pilots' proposal to purchase United does
           not bespeak a lack of confidence in Mr. Ferris.

2.2.3 Adjuncts

The segmentation of adjuncts is contingent upon whether or not they encode an
eventuality, i.e. some sort of event or state. As a result, purely temporal or locative
adverbials are not treated as EDUs.

Clausal Adjuncts   Clauses introduced by subordinating conjunctions such as *when*,
while, etc., are treated as EDUs. This follows the event rule stated above, as such
conjunctions generally take tensed clauses, which normally introduce eventualities,
as arguments.

(6)   For the second time in a week, a widely recognized
       American figure has died<***> when he was skiing and
       ran into a tree.

(7)   He was serving his second term in Congress<***> when
       he died late yesterday at a ski resort on the border...

Infinitival clauses functioning as purpose clauses are treated as separate dis-
course units. Purpose clauses are identified with the *in order to* test: replace *to*
with *in order to* and check for semantic equivalence. For example:

(8)   U.S. troops acted for the first time to capture an
       alleged Bosnian war criminal,<***> rushing from unmarked
       vans parked in the northern Serb-dominated city of
       Bijeljina<***> to seize a former concentration camp
       commander accused of killing at least 16 Muslims and
       abusing or terrorizing scores of others.

When the 'in order to' test does not give a clear answer as to whether one should
segment, another rule of thumb is that there must be an agent acting in order to do

5

something when we segment 'to' phrases. If the 'to' phrase merely describes the purpose of a project or plan, for example, then we do not segment.

(9)   `The Justice Department's lengthy review of the deal`
        `is already threatening to scuttle Primestar's `<u>`plans to launch a sat`</u>
        `this summer.`

Other Adverbials   The segmentation of other adverbials, especially adverbial PPs, is more problematic. Again, annotators should obey the event rule. For example, adverbial PPs are segmented only if they contain a nominalized **event**, such as *after the meeting*. Frame adverbials, which denote a **time** (e.g. 'Last Sunday...') or **location** (e.g. 'In Jakarta...'), are not treated as discourse segments.

    Following the event rule, *by*-phrases introduce discourse segments, as in (10).

(10)   `The UAL board is four-square behind Mr.  Ferris, his`
        `management team and his long-range strategy of making`
        `United a more competetive airline<***> by combining`
        `it with the premier hotel company and car rental company.`

As phrases call for special comment. We have segmented such adverbials when fronted. Consider for instance

(11)   `As vice premier and president of the People's Bank`
        `of China in 1995,`

(12)   `As mayor,`

These do describe activities that are temporally bounded. Are these events? It's relatively plausible, though we do see a need for giving a definition of event and eventuality. Davidson individuates events in terms of causes and effects (and defines causes and effects in terms of events). Aside from its circularity, Davidson's definition is of little use to us. As far as we are aware there are no definitions of events in the Neo Davidsonian literature. Kim, Lemmon and Lewis each give precise and non circular definitions of events. On Kim's view events are closer to facts (they are spatio-temporally located complexes of instantiated properties), while on Lemmon's view they are simply regions of space time and on Lewis's they are functions from worlds to regions of space time. None of these views are of any use in making our event criterion more precise.

2.2.4 Appositives

Appositives are always segmented unless segmentation would result in a discontinuous EDU. Note that not all appositives conform to the event rule for adjuncts detailed above.

(13) "I look forward to their return on March 9,"<***> said
      Irish Foreign Minister David Andrews,<***> who with
      Mowlam met with Adams and other Sinn Fein leaders in
      Belfast today.

(14) The other two slots are being used by Primestar's competitors,
      <***> DirecTV Inc. and Echostar Communications Corp.

(15) but that's certainly not the case here<***> said Bob
      Dornan,<***> senior vice president of Federal Sources,<***>
      a McLean-based consulting firm.

2.2.5 Coordination

Coordinated tensed clauses function as discourse units; however, nonclausal coordinate structures, such as VPs, introduce EDUs only in certain cases. Coordinated VPs are treated as separate discourse segments when they either include a discourse particle or contain discourse structure within (at least one of) the coordinated constituents. In (16), the conjoined NPs are segmented because the second conjunct contains the discourse marker *then*.

(16) Congressman Sonny Bono was first a songwriter in the
      nineteen sixties<***> and then a popular entertainer.

3 Discourse Relations

In this section, we discuss the rhetorical relations used to annotate texts. For each relation, we describe its semantic effect on the interpretation of a text and the kind of surface cues that are indicative of the relation. We assume that semantic effect and surface linguistic form together provide sufficient evidence that two discourse units are related by a given relation. All 14 discourse relations that we used are listed in Table 1.

SDRT groups each rhetorical relation according to the structural configuration that it yields in a discourse graph. There are two such configurations, those provided by coordinating relations and those provided by subordinating relations.

Coordinating Relations		Subordinating Relations	
Vericical	*Nonveridical*	*Veridical*	*Nonveridical*
Continuation	Consequence	Background	Attribution
Narration	Alternation	Elaboration	
Result		Explanation	
Contrast		Commentary	
Parallel		Source	
Precondition			

Table 1: Discourse Relations used in the Annotation Task

SDRT also distinguishes **veridical** from **nonveridical** relations. The classification of discourse relations according to these factors is shown in Table 1. Veridical relations entail the content of (both of) their arguments, whereas non-verdical relations fail to entail the content of at least one of their arguments.

3.1 Subordinating Relations

We begin by describing the subordinating rhetorical relations listed in Table 1, using α and β as variables for EDUs.

3.1.1 Veridical Relations

Background $Background(\alpha, \beta)$ holds when a constituent β provides extra information about α. One consequence of $Background$ (and a cue for it) is that when it obtains between two EDUs that describe eventualities there is temporal overlap between the two related eventualities. In such cases, $Background$ is often signaled by aspectual shift, i.e., a shift from an event to a state, or state to an event. Clauses introduced by subordinating conjunctions such as *when* and *while* also are cues for $Background$ to the EDU given by the matrix clause. With respect to *while* clauses, the generalization usually holds when the subordinate clause is preposed; however, some care is required in the case of postposted *while* clauses, as they sometimes give rise to $Contrast$ or $Explanation$ (see below). Sometimes the Background relation is also reversed and the postposed clause is the foreground while the main clause is the background, as in ((18) below.

Example (17) shows a typical use of $Background$. $Background$ is signaled here by an apsectual shift between the related discourse units: the first sentence describes a past **event** and the second introduces a **generalizing stative**.

(17)   a.   Also, about 585 workers were laid off at a stamping
               plant near Detroit.
         b.   That plant normally employs 2,800 hourly workers.

(18)   a.   He had been on duty 15-20 minutes,
         b.   when Lewinsky arrived saying she had some paperwork
               she needed to bring into the President.

We have found that *Background* also has uses when no eventualities are explicitly involved. *Background* often links appositive phrases to other EDUs; nonrestrictive, or appositive, relative clauses often attach with *Background*, and nominal appositives always do. In these cases, we take Background to be a way of furnishing more information about one or more objects in the foreground EDU. In

(19)   said Zhang Xiuxue,*** vice president of the Chinese
         National School of Administration.

(20)   As vice premier and president of the People's Bank
         of China in 1995,*** Zhu clamped down on the free flow
         of credit from provincial branches of the bank to projects
         and state-owned enterprises favored by provincial officials.

(21)   and the government began considering establishing a
         ``currency board" that would rigidly fix the value
         of the rupiah against the U.S. dollar --*** a move
         Washington dismissed as unworkable.

(22)   JAKARTA, Indonesia, Feb. 27---*** James Riady, the
         Indonesian businessman who befriended President Clinton
         in Arkansas and became a major contributor to the Democratic
         Party, today denied allegations...

(23)   Despite his dreams,*** Riady still has to deal with
         the short term

(24)   that it would buy Franklin Bancorp Inc., *** a small
         D.C. bank.

(25)   while Maryland Federal's stock closed at $35.75, ***up$
         4, on the Nasdaq Stock Market.

Elaboration   *Elaboration*(α, β) holds when β provides further information about the eventuality introduced in α; for example, if the main eventuality of β is a sub-

type or part of the eventuality mentioned in α.[1] When *Elaboration* holds between two eventualities, it implies that a relation of temporal inclusion holds between the related eventualities. Discourse markers like *for instance*, *for example*, or the explicit listing of sub-events (*first*, *second*, etc.), are good cues for *Elaboration*. RST includes a number of specialized elaboration relations, which we subsume under our *Elaboration*.

By phrases can also sometimes indicate Elaboration, specifically when they specify a manner in which some eventuality is carried out. An example is given in ((26)):

(26)   But European allies predict<***> that Milosevic will
       respond to encouraging gestures<***> by taking further
       steps toward peace.

(27) provides an example of *Elaboration* in which the second EDU provides more information about the deadlock mentioned in the first EDU.

(27)   a.   Both sides were deadlocked.
       b.   Areas of dispute include use of temporary workers
            at NBC, the length of the new contract and use
            of non-network news services.

The second EDU in (28) precisifies the point made in the first EDU. This precisification is signaled by 'namely', another cue word for *Elaboration*.

(28)   Riady still has to deal with the short term,<***> namely,
       wrestling with Indonesia's economic crisis.

In the following example, the disjoined adverbials precisify the state of being alone with Lewinsky.

(29)   that he does not recall <***> ever being alone with
       Lewinsky,<***>either while she was employed at the
       White House or later at the Pentagon

Elaboration can hold in cases of event restatement—i.e., when the elaborating clause re-describes the event in the elaborated clause. In the following example, ((30-b)), together with ((30-c)), restates the event described in ((30-a))

---

[1]Eventualities can be introduced by nominal expressions, viz. those that denote eventualities in some way.

(30) a. Albright addressed this perceptual problem at the
          December meeting,
      b. telling her European colleagues
      c. that "too often, the United States takes the heat
          for dealing with difficult issues

Not all elaborations hold between EDUs that introduce, in dynamic semantic terms, two eventualities (i.e., have an existential quantification over eventualities). It would be very surprising, if this were true. In ((31)) for example the first EDU contains a quantification over eventualities (i.e., "Not every conversion to state owned corporation is guaranteed of success") but it does not describe any particular eventuality.

(31) Conversion to state-owned corporations is no guarantee
      of improved performance, however.<***>The petroleum
      ministry was converted to the Chinese National Petroleum
      Corp. several years ago...

Nevertheless, the second EDU in (31) provides an example to support the point made in the first EDU. Thus, the spatio temporal consequences for Elaboration should only hold when the discourse constituents linked introduce eventualities.

Similarly ((32)b) below doesn't describe an eventuality. This situation is already considered in (Asher 1993), where *Elaboration* also covers cases in which there is a relation of defeasible implication between constituents. Thus if A defeasibly implies B, then linking B to A is a way of elaborating on A (SDRT also includes providing evidence as a kind of elaboration). This implicational link is at work in ((32). An implicational link from B to A also serves to indicate *Elaboration*. This holds for instance in ((31)).

(32) a. Carl is a tenacious fellow.
      b. He doesn't give up easily.

Sometimes *Elaboration* cue words are followed by noun phrases only, as in the following two examples. In these cases, we still use *Elaboration* to relate the two EDUS.

(33) only the president's assistants allow people in to
      see him,<***>even on weekends.
      (34) Starr has refrained from subpoenaing any Secret
           Service officers,<***>including Fox.

In annotation practice, we have had trouble distinguishing between *Elaboration* and *Background* in some cases. In particular, the problem revolves around constituents with semantically light verbs—verbs like *have, contain, is, detail* and a host of others. We need to make a complete catalogue of these. The constituents with light verbs don't seem to the annotators to give much of an eventuality, and so the intuition is that Elaborations in these cases can hold in virtue of relations between eventualities in the elaboration and in some argument of the light verb in the topic constituent, or constituent that is elaborated.

Explanation   When α and β introduce eventualities in the dynamic sense (i.e. existential quantification over eventualities occurs with wide scope over modal operators, negation or non-existential quantifiers), $Explanation(\alpha, \beta)$ holds when the main eventuality of β is understood as the cause of the eventuality in α. *Explanation* has temporal consequences, viz. that the eventuality described in β precedes (or overlaps) the eventuality described by α. *Because* is a monotonic cue for *Explanation* (see (35)).

(35)   a.   The department last week rejected TWA's first application as "deficient"
           b.   because it omitted such important information as the merger's potential impact on competition and pricing.

'After' and 'when' sometimes signal *Explanation*, as in the following two examples.

(36)   BTG Inc.'s chairman yesterday rescinded the dismissals of four vice presidents who were unexpectedly fired Monday<***>after an abortive bid to purchase a BTG division'' hidden explanation

(37)   14 Palestinians were injured Saturday<***> when Israeli troops fired rubber-coated metal bullets at a few hundred demonstrators who attacked them with stones.

Purpose clauses are always related via *Explanation* to the matrix clause, as in (38). In this case, the relation receives an intentional interpretation. In such cases the temporal consequences of Explanation do not follow because the eventuality introduced under the scope of an intensional operator. For example in ((38)b) is understood as roughly equivalent to *because they wanted to put pressure on swing*

votes in the Senate and House.[2]

(38)   a.   `Both parties are using business leaders as proxies of sorts`

           b.   `to put pressure on swing votes in the Senate and House,`

In (38), both parties are using business leaders as proxies not only to put pressure on swing votes in the Senate and House, but also because they *want to* put pressure on swing votes in the house and senate. This use of *Explanation*, therefore, is nonveridical.

In some cases, when two EDUS are related by *Explanation*, the explanans explains how, as opposed to why, the explanandum is the case. 'By' phrases can signal such uses of *Explanation*.

(39)   a.   `They were able to raise the necessary funds`

           b.   `by cashing in all their stock options.`

In these cases, however, there should be a paraphrase of the two constituents in which we have 'because'.

(40)   a.   `John made Mary angry`

           b.   `by kissing her.` That is in ((40)) John made Mary angry <u>because</u> he kissed her. Similarly for ((39), they were able to raise the necessary funds, <u>because</u> they cashed in all their stock options.

Source and Attribution $Source(\beta, \alpha)$ and $Attribution(\alpha, \beta)$ are used to relate the content of a communicative act, given in β, to the agent of that act, given in α. Both relations are subordinating, but *Attribution* is intensional, and consequently nonveridical on the right, which means that what is said cannot be taken as fact. On the other hand, *Source* is veridical in both arguments; consequently with *Source* what is communicated is asserted as fact. *Source* and *Attribution* are structurally distinguished as well. When $Source(\beta, \alpha)$ holds, the matrix clause, α, is subordinate to the embedded clause, β; whereas when $Attribution(\alpha, \beta)$ holds, these relationships are reversed: the embedded clause β is discourse-structurally subordinate to the matrix clause α.

Source is the default relation used to connect communicative agents to the content of their communicative acts. Semantically, *Source* marks a kind of evidential use of communicative verbs. Thus, *Source* should be used when:

---

[2]p.c., James Pustejovsky.

- the content of the communicative act carries the main rhetorical import;

- the communicative agent is in "a position to know", e.g. when agents discuss their own mental states, or a company spokesperson talks about the actions of the company;

- the verb of communication occurs as a parenthetical expression; and

- lexical items that presuppose the truth of their complement such as *acknowledge, admit, confirm*, *clearly* or *as previously reported* occur in the matrix clause.

The example in (41) illustrates the first two bullet points. First, the main rhetorical import of the complex segment is that NBC will put its contract offer into effect. Second, the agent of the communication, viz. NBC, is in a position to know about the actions of the network.

(41)   a.   `National Broadcasting Co.  told its technicians`
                `and news employees union`
       b.   `that the network will put into effect next Monday`
                `its latest contract offer.`

(42) illustrates the third bullet point. In this example the communicative verb occurs in the sentence final parenthetical expression *a source close to the board said*.

(42)   a.   `But Pan Am, sensing progress in its recent negotiations`
                `with the unions, asked the board to delay acting`
                `on the company's request,`
       b.   `a source close to the board said.`

Finally, because the verb *acknowledged* is factive, the main rhetorical import of the segment is assumed to be given by the embedded clause. Because of this (43-a) is linked to (43-b) with *Source*.

(43)   a.   `Many present and former officials in the middle`
                `and lower ranks acknowledged privately`
       b.   `that they did not see Clinton's careful statements`
                `yesterday as anything like the full-throated denial`
                `they were hoping for.`

Attribution relates a communicative agent and the content of a communicative act when this content is not taken to form part of the story line attributable to the author of the text. Therefore, *Attribution* is used when:

- the main rhetorical import of the segment is that an agent x has said ϕ;

- two attributions provide contrasting viewpoints or contradictory allegations;

- the embedded clause contains evaluative adjectives, verbs or abverbs;

- lexical items such as *insisted*, *reportedly* or *complained* are used.

(44) demonstrates the first two bullet points. Here the main rhetorical effect of each attribution is that someone made a claim and, in some sense, these claims are incompatible or contradictory.

(44)   a.   `The union said`
        b.   `the size of the adjustments was inadequate.`
        c.   `But Chrysler Canada's chief negotiator, William`
            `Fisher, said yesterday`
        d.   `that the two sides had reached "some understandings"`
            `on economic issues, including pensions.`

Example (45) demonstrates the third bullet point, as the agent is voicing his opinion about something already under discussion in the story.

(45)   a.   `It's a serious application,`
        b.   `a department official said of the new filing.`

Example (46) illustrates the fourth bullet point. Certain lexical items, like *reportedly*, indicate that the main rhetorical effect is the communicative act itself, rather than the content of the communication.

(46)   a.   `Mr.  Icahn reportedly said`
        b.   `he "couldn't watch that happen."`

In practice, we have found it sometimes difficult to decide between *Attribution* and *Source*. We believe that the placement of the attributive verb as a parenthetical or after the subordinate clause is an important clue indicating *Source*. But it is not decisive. We have several examples in which structural considerations—e.g., we have a contrast between two *Source* or two *Attribution* relations—play an important role. This issue needs further clarification and analysis for our annotations to become more consistent with each other.

Commentary   *Commentary*(α, β) holds if β provides an opinion or evaluation of the content associated with α (Asher 1993). "Supplementary adverbs" (Potts 2005) are good surface cues for *Commentary*. These include speaker-oriented

adverbs, such as supplemental uses of "luckily", "amazingly", etc, and utterance modifiers such "frankly", "confidentially", etc. The last EDU in (47), for example, will attach to the discourse context via *Commentary* given the presence of the utterance modifier *frankly*.

(47)　a.　"Yesterday's performance was a departure,"
　　　b.　Mr. Callahan said,
　　　c.　[...]
　　　d.　"Frankly, it's a little mystifying."

(47) also provides an example of the interaction of *Attribution* and *Source* that is often found in the corpus. Mr. Callahan's communicative act will attach to the preceding discourse with *Attribution*, because it does not constitute part of the main storyline attributable to the author, but rather provides one agent's commentary or evaluation of the events introduced in the preceding text, viz. that they are mystifying.

Precondition　*Precondition*(α, β) is used to represent "anti-narration" (see below). That is, *Precondition* is like *Narration* except that the order of the arguments is reversed. *After* clauses often give rise to *Precondition*, as in (48).

(48)　a.　Mr. Gerstner raced to hire Mr. York
　　　b.　after meeting him for the first time just three
　　　　　weeks ago in IBM's Manhattan offices.

Here the temporal trace of the main eventuality of β precedes that of the main eventuality described in α. Paraphrasing, Mr. Gerstner first met Mr. York three weeks ago, *then* he raced to hire him.

　　Commentary and *Precondition* are not that frequent in our corpus.

3.2　Coordinating Relations

3.2.1　Veridical Relations

Narration　*Narration*(α, β) holds when the main eventualities of the EDUs α and β occur in sequence and have a common topic. Certain spatio-temporal consequences follow from *Narration*; for example, the pre-state of the eventuality associated with β overlaps with the post-state of the eventuality associated with α. *And* and *then* are good monotonic cues for *Narration*.

(49)　a.　it expects to borrow as much as $2 billion from
　　　　　banks to complete the acquisition.

b. It then plans to refinance the bank loans with
proceeds from public or private equity and subordinated
debt offerings.

Continuation $Continuation(\alpha, \beta)$ is like *Narration* without the spatio-temporal consequences. *Continuation* often holds between two EDUS when they both elaborate or provide background to the same EDU. Nevertheless, it is not sufficient that both edus provide, e.g., background information. They also have to be suitably related. For instance, if one constituent β is linked with *Background* to α because it provides more information about an object a while another constituent γ is linked with *Background* to α because it provides more information about the time the main eventuality in α, then we will not add $Continuation(\beta, \gamma)$ to the discourse structure. In the following example, both (((50-a)) and (((50-c)) are related to (((50-b)) via *Background*, but because ((50-a)) and ((50-c)) do not form one continuous thought, we do not relate them via *Continuation*.

(50)   a. So long as the specter of violence continues to
loom,
       b. their lack of confidence in Indonesia may persist,
       c. regardless of what reform measures Suharto embraces.

In (51), in contrast, (51-b) and (51-c) are related to (51-a) via *Background* and to each other via *Continuation*. They are suitably related because they both give more information about talks between American officials and the employees of the airline.

(51)   a. American officials "felt talks had reached a point
where mediation would be helpful."
       b. Negotiations with the pilots have been going on
for 11 months
       c. ; talks with flight attendants began six months
ago.

Of all the coordinating relations *Continuation* is perhaps the most frequently used, and it often occurs with other relations like *Narration*, *Result* or *Contrast*.

Contrast $Contrast(\alpha, \beta)$ holds when α and β have similar semantic structures, but contrasting themes, i.e. sentence topics, or when one constituent negates a default consequence of the other. *But, however, on the other hand, nevertheless* are all strong cues for *Contrast*. *While*-clauses also sometimes introduce *Contrast*, whether they are in postposed or preposed position.

(52) a. And VHS-C can be played on existing VHS machines
          with only a simple adapter;
     b. by contrast, 8mm tapes are incompatible with the
          VHS machines.

(53) a. While European allies have closed ranks behind
          the United States in the latest showdown with Iraqi
          President Saddam Hussein,
     b. there are lingering apprehensions among NATO governments
          about the Clinton administration's insistence on
          recognizing the spread of nuclear, biological and
          chemical weapons as the alliance's most urgent
          priority.

With many of the syntactically subordinating discourse connectors like the adverbials 'while', 'except', 'unlike'. Not 'despite', annotators have felt a conflict between annotating these with a subordinating relation like *Background* but also with *Contrast*. Annotating with *Background* and *Contrast* is fine, if we take the position of **?** on which structural relations like *Contrast* are neither subordinating nor coordinating and so are compatible with either.

Parallel   *Parallel*(α, β) has the same structural requirements as *Contrast*, but instead requires α and β to share a common theme. Cue phrases such as *too* and *also* are good indicators of *Parallel*.

(54) a. The United Telegraph Workers union represents 4,400
          Western Union employees around the country.
     b. The Communications Workers of America represents
          300 company employees in New York City.

(55) a. Zhu clamped down on the free flow of credit from
          provincial branches of the bank to projects and
          state-owned enterprises favored by provincial officials.
     b. He also overhauled tax revenue sharing

Result.   *Result*(α, β) relates a cause to its effect: the main eventuality of α is understood to cause the eventuality given by β. Thus *Result* is the dual of *Explanation*, which relates an effect to its cause. (56) and (57) provide illustrations. Note the cue phrase *as a result* and the discourse marker *so*.

(56) a. Chrysler stopped output of Dodge Dynasty and Chrysler
          New Yorker cars yesterday at its Belvidere, Ill.,

```
        plant
    b.  Some 1,700 of the plant's 2,900 hourly employees
        were laid off as a result.
(57)    a.  The mediator has imposed a news blackout on the
        two sides,
    b.  so a Big Board spokesman couldn't comment on the
        talks.
```

3.2.2 Nonveridical Relations

Consequence and Alternation   *Consequence* and *Alternation* are nonveridical, coordinating relations and correspond semantically to the logical operators $\rightarrow$ and $\vee$ respectively. *Consequence* is normally introduced by *if...then* statements or certain quantificational statements, as in the first two examples below, while *Alternation* is normally triggered by *or*, as in the third example below.

```
(58)    a.  If a settlement comes today or tomorrow,
    b.  Chrysler's 10,000 Canadian workers could vote on
        a contract this weekend,

(59)    a.  When they do make an appearance,
    b.  it is generally to express blandly their support
        for the Iraqi people and their hopes for a diplomatic
        solution and to complain that Washington is treating
        Baghdad unfairly.

(60)    a.  either by TWA's acquisition of USAir,
    b.  or USAir's acquisition of TWA.
```

The example below is also treated as *Alternation*, which follows from the logical equivalence of $\phi \vee \psi$ and $\neg\phi \rightarrow \psi$.

```
(61)    a.  "It looks like we're going to be there on the street,
    b.  unless there is a miracle,"
```

We abuse *Alternation* slightly by using it to account for examples like (62). Intuitively in (62), the state described in the first EDU will hold until the state or event described by the second EDU obtains.

```
(62)    a.  Several appointees of President Bush are likely
        to stay in office at least temporarily,
    b.  until permanent successors can be named.
```

Paraphrasing loosely, either ϕ or ψ holds, but not both ϕ and ψ—just as in exclusive disjunction. For this reason, we annotate such examples with *Alternation*. However, this choice is semantically incorrect, as whatever relation in fact relates the two EDUs in (62) should entail the truth of the left argument, viz. that Bush's appointees will stay in office.

4 The Annotation of Discourse Structure

4.1 SDRT **Basics**

The annotation scheme for encoding rhetorical structure is a straight forward encoding of segmented discourse representation structures, or SDRSs. An SDRS represents the rhetorical connections between various segments of a text and the hierarchical structure of discourse. Formally, an SDRS is a tuple $\langle A, \mathcal{F} \rangle$, where

- A is a set of labels, and

- $\mathcal{F}$ is a function from A to the set of well-formed SDRS-formulae.

The well-formed SDRS formulae include the logical forms for elementary discourse units; specifically, they include single clauses, formulae of the form $R(\alpha, \beta)$ (where R is a rhetorical relation and α and β are labels in A), and the dynamic conjunction of SDRS formulae. The labels in A include labels for elementary discourse units, as well as labels for larger text segments.

The constructed discourse in (63) illustrates many of the important features of SDRSs.

(63)   a.   John had a lovely evening last night. (π_1)
       b.   He had a great meal. (π_2)
       c.   He ate salmon. (π_3)
       d.   He devoured lots of cheese. (π_4)
       e.   He then won a dancing competition. (π_5)

Intuitively, the EDUs (63-b) and (63-e) elaborate the main eventuality described in (63-a), namely John's lovely evening. (63-e) continues a narrative segment begun by (63-b). First, John had a great meal. Then he won a dancing competition. The segments given by (63-c) and (63-d) elaborate the great meal and together form a second narrative segment.

The SDRS $\langle A, \mathcal{F} \rangle$ for (63) is provided below.

- $A = \{\pi_0, \pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6, \pi_7\}$

- $\mathcal{F}(\pi_1) = K_{\pi_1}$
  $\mathcal{F}(\pi_2) = K_{\pi_2}$
  $\mathcal{F}(\pi_3) = K_{\pi_3}$
  $\mathcal{F}(\pi_4) = K_{\pi_4}$
  $\mathcal{F}(\pi_5) = K_{\pi_5}$
  $\mathcal{F}(\pi_0) = Elaboration(\pi_1, \pi_6)$
  $\mathcal{F}(\pi_6) = Narration(\pi_2, \pi_5) \wedge Elaboration(\pi_2, \pi_7)$
  $\mathcal{F}(\pi_7) = Narration(\pi_3, \pi_4)$

The labels π_6 and π_7 represent spans of text corresponding to two or more EDUs. π_7, for example, represents the narrative segment composed of the EDUs π_3 and π_4 which elaborate the great meal introduced in π_2.

An SDRS $\langle A, \mathcal{F} \rangle$ can be visualized as a directed acyclic graph in which the labels in A provide vertices, or nodes, and the rhetorical connections between labels introduce labeled edges between nodes. If an SDRS contains the formula $R(\alpha, \beta)$, then the corresponding graph contains an edge from the node α to the node β labeled R. If R is a coordinating relation, the edge is horizontal; if R is a subordinating relation, the edge is vertical. The graph encoding of (63) is provided in Fig. 1. The formula corresponding to the top most label in the SDRS for (63), which is omitted in the graph, is $Elaboration(\pi_1, \pi_6)$. Since $Elaboration$ is a subordinating relation, the graph for (63) contains a vertical edge from π_1 to π_6 labeled $Elaboration$.
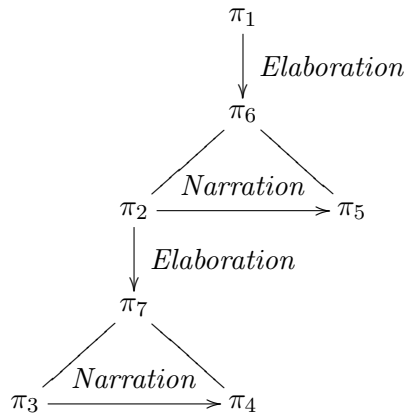

Figure 1: SDRS graph for (63).

During the construction of an SDRS only a subset of the labels in A are available for attachment, namely those on the right-frontier of the graph. In the graph in

figure 1, for example, only π_1, π_6 and π_5 are for the attachment of additional material. As a result, (64) could not be used to elaborate on the salmon introduced in π_3.

(64)   #It was a beautiful pink.

4.2   The Two-Fold Path to Infinite Wisdom

We have tried several ways of segmenting and then labelling texts. The extreme complexity of the task makes it difficult to get strong inter annotator agreement on the finished product, when annotators begin just from the segmented text. We have now come up with a two stage process that produces better results. In the first stage, annotators group segments together; segments that are grouped together will be attached to each other. At this stage no discourse relations are entered into the annotation. The stage is purely about attachment and subordinating and co-ordinating information. Once the groups have been distinguished, annotators can then think about how groups attach to the rest of the story. We have found typographical (paragraph) information very important in constructing groups, though not altogether decisive.

Helpful in the following cases: one edu looks like a Background on a second edu. The first edu is clearly coordinated with a third edu by a relation such as Contrast, yet the third edu should be attached to the second via Explanation. In this case, the contrast between the first and the third edus may be stronger than the relation background relation between the first and second; thus we will attach the first to the second via Explanation. The grouping technique gives us a more holistic picture. If we approach the task linearly, then we get a different interpretation. For example, in the example given above, if we would have attached the first edu before looking at its relation to the third, it would have been attached via Background. We found these different approaches to be one source of inter-annotator disagreement.

4.3   The Current Annotation Scheme

In the current annotation scheme the graph for the text in (63) shown in Fig. 1 is represented as follows:

- `Elaboration(1,[2,5])`
  `Elaboration(2,[3,4])`
  `Narration(3,4)`
  `Narration(2,5)`

Each line consists of a relational formula stating the rhetorical connection between two discourse units. Elementary discourse units are assigned a positive integer as a label, and larger discourse units are assigned a unique identifier based on the segments that they immediately dominate in the SDRS graph representation. π_6 from Fig. 1, for example, is assigned the identifier `[2,5]` because it immediately dominates π_2 and π_5. [3] The identifiers for the subsegments are placed in square-brackets and separated by commas. We take such complex identifiers to be names of complex discourse units, comparable to π_6 and π_7 from Fig. 1.

The construction of non-elementary discourse units becomes especially relevant when a series of contiguous discourse units are connected by coordinating discourse relations. In such cases potential ambiguities arise with regard to the interpretation of the discourse. Consider the discourse in (65) from Asher (1993). On one interpretation, π_4 relates to a complex discourse unit composed of π_1, π_2 and π_3 via *Contrast*. On this reading, the anaphoric pronoun *this* in (65-d) refers the summation of the claims made by the three plaintiffs.

(65)   a.   One plaintiff was passed over for promotion three times. (π_1)
       b.   Another didn't get a raise for five years. (π_2)
       c.   A third plaintiff was given lower wage compared to males who were doing the same work. (π_3)
       d.   But the jury didn't believe *this*. (π_4) This reading is encoded in our annotation scheme as follows:

- ```
  Continuation(1,2)
  Continuation(2,3)
  Contrast([1,2,3],4)
  ```

On another reading, however, π_4 is connected to π_3 with *Contrast*, and the pronoun refers only to the claim made by the third plaintiff.

In general, any situation in which three or more discourse units are related by coordinating discourse relations is three ways ambiguous. The general situation is schematized as follows:

$$\alpha \xrightarrow{R_1} \beta \xrightarrow{R_2} \gamma$$

Given such a situation an annotator could potentially encode in one of three ways. First:

---

[3] π *immediately dominates* π' in an SDRS graph iff either (i) $\mathcal{F}(\pi)$ contains as a conjunct $R(\pi', \pi'')$ or (ii) $R(\pi'', \pi')$ and *coordinating*(R).

- `R_1(a,b)`
  `R_2(b,c)`

On this reading, β relates to α through R_1 and γ relates to β through R_2. This roughly corresponds to the second interpretation of (65) above. (66) illustrates.

(66)   a.   In the Ford-UAW talks, the two parties turned to economics yesterday in addition to the job-security issue. (54)

         b.   [...]

         c.   On the job-security issue, it is understood that Ford agreed not to lay off any workers during the life of the contract except during a major sales plunge. (59)

         d.   The company also promised to replace half of all workers leaving the company. (60)

         e.   At last report, the two sides were still haggling over details of the job-security plan, (61)

In (66) the complex segment `[60,61,62]` elaborates `54` and each EDU in `[60,61,62]` is connected by *Continuation*. Thus (66) is annotated as follows:

- `Elaboraton(54,[60,61,62])`
  `Conintuation(60,61)`
  `Continuation(61,62)`

Alternatively, the annotator might decide that β and γ are related by R_2 and that α is related to the complex segment formed by β and γ. This interpretation is encoded as follows:

- `R_1(a,[b,c])`
  `R_2(b,c)`

(67) provides an example of this situation from our corpus. The EDUs `10` and `11` are related with *Contrast*. This complex segment continues EDU `10`.

(67)   a.   the Guild, which represents nearly 300 of the Post's 700 employees, voted Saturday to end its strike. (10)

         b.   Guild members have been invited to reapply for their old jobs today, (11)

c. though it is assumed that dozens of them won't
be rehired. (12)

Finally, α and β can potentially form a complex segment related to γ by the discourse relation R_2. Schematically, this situation would be annotated as follows:

- R_1(a,b)
  R_2([a,b],c)

The text segment in (68) illustrates this situation. The EDUs in 65 and 66 are related by *Contrast*, as indicated by the discourse cue phrase *by contrast*, and EDU 67 continues this complex segment.

(68) a. Chrysler last year paid Mr. York $414,167 in salary,
          a $300,000 bonus and $327,621 in long-term compensation
          based on quality improvements in the company's
          products. (65)
      b. By contrast, IBM paid Mr. Metz, its former chief
          financial officer, a total of $725,000 last year.
          (66)
      c. Mr. York also had stock options that Chrysler
          valued at $2.9 million at the end of 1992. (67)

Consequently, (68) is the mirror image of (67).

4.4 Other Issues

Importantly, SDRSs represent directed acyclic graphs, not trees. The non-treeness of discourse structure assumed by SDRT manifests itself in several ways, e.g. multiple relations holding between two utterances and multiple parents. Since our annotation scheme is a straightforward encoding of a relation structure, we have no problems handling such cases if/when they arise. We refer the reader to Baldridge and Lascarides (2005) for a more thorough discussion of these issues.

Our annotation scheme sometimes leads us to reverse EDUS. For instance, $Consequence(\alpha, \beta)$ encodes a conditional in its "canonical" form—$if\,p, then\,q$. However, our corpus is replete with examples in which the antecedent follows the consequent—viz., we have $q, if\,p$, which we encode as follows, where p is constituent 1 and q is constituent 2: $Consequence(2, 1)$. We annotated the following text as: Elaboration(a,[c,b]), Consequence(c,b).

(69) a. by threatening farmers
      b. with sanctions

```
        c.  if they don't limit the amount of manure they use
            as fertilizer.
```

Preposed adverbial clauses or adverbial phrases that are segmented and attach with *Background* will also lead to formulas of the form $Background(m + n, m)$.

```
(70)    a.  So, with television cameras rolling and everyone
            dripping wet,
        b.  they resorted to tearing the flag to pieces.
```

The correct annotation of this text is Background(b,a).

References

Asher, N.: 1993, *Reference to Abstract Objects in Discourse*, number 50 in *Studies in Linguistics and Philosophy*, Kluwer, Dordrecht.

Asher, N. and Lascarides, A.: 2003, *Logics of Conversation*, Cambridge University Press.

Baldridge, J. and Lascarides, A.: 2005, Annotating discourse structures for robust semantic interpretation, *Proceedings of the Sixth International Workshop on Computational Semantics (IWCS)*, Tilburg, The Netherlands.

Carlson, L., Marcu, D. and Okurowski, M. E.: 2003, Building a discourse-tagged corpus in the framework of rhetorical structure theory, *in* J. van Kuppevelt and R. Smith (eds), *Current Directions in Discourse and Dialogue*, Kluwer Academic Publishers, pp. 85–112.

Potts, C.: 2005, *The Logic of Conventional Implicatures*, Oxford University Press.

The PDTB Research Group: 2006, The Penn discourse treebank 1.0 annotation manual, *Technical report*, Institute for Research in Cognitive Science, University of Pennsylvania.

Wolf, F. and Gibson, E.: 2005, Representing discourse coherence: A corpus based study, *Computational Linguistics* **31**(2), 249–287.