Annotating Argumentation Acts in Spoken Dialog

Amanda J. Stent and James F. Allen

The University of Rochester
Computer Science Department
Rochester, New York 14627

Technical Report 740 and TRAINS Technical Note 00-1

December 2000

Abstract

This manual describes a method for annotating rhetorical relations, adjacency pairs, and other argumentation acts found in task-oriented spoken dialog [9, 10]. It is largely aimed at the novice annotator rather than the computational linguist, and therefore in choosing terminology we have valued intuitiveness over precision. This work came out of an exploration of how to mark structure above the speech act in the Monroe corpus [7]. For more information about the development of this manual, see [8].

This tool is designed for use with ArgumentationTool, a tool for marking argumentation acts in dialog that is available from `http://www.cs.rochester.edu/research/cisd/resources/aad/`.

This work was funded by DARPA research grant no. F30602-98-2-0133 and ONR research grant no. N00014-95-1-1088. Our thanks go to Jenine Turner, Christi-Lynn Brown and Michael Jackson, our initial annotators, who provided many valuable comments.

Contents

1 Introduction

You will be annotating a set of dialogs. Annotation is a process of marking useful information. Sometimes text or speech are annotated for things like the parts of speech of words; sentence or document structure; or the names of people or locations.

You will be marking the "why" of parts of dialog, the relationships that identify the reasons people say sequences of things. That is, you will mark the patterns of speaking that people use to convince, to motivate, to inform and so on. We will call these patterns **relations**.

In order to do this marking, you will need to know the answers to several questions:

- How is a dialog broken into parts? What is the size of one dialog "chunk"? (section 2)
- How will I know when I can stop? (section 3)
- How do I mark relations? (section 4)
- What does the annotation process look like? (section 5)
- What are the specific relations I should consider? (section 6)

This manual answers these questions.

1.1 Before starting

It is important to read the whole manual before doing any annotation, and to follow the annotation instructions carefully. The annotator's task is difficult, but with practice and attention you will find it becomes easier over time.

1.2 Getting ready

In addition to this manual, you will need several other things to be able to do your annotation:

1. An annotation tool. The tool this manual was written for, ArgumentationTool, is available from `http://www.cs.rochester.edu/research/cisd/resources/aad/`, but other tools may also work. Information about running ArgumentationTool is available at the back of this manual.

2. A text editor. As you annotate, you will make comments in a comments file. The annotation tool provides a way for you to make comments, but if you use another tool or want more space to see your comments, you may use any text editor (e.g. Emacs, Word, the Windows Notepad).

3. Dialogs to annotate. A selection of dialogs is available at the same address as the tool. Whatever dialogs you are annotating, you will want the following for each dialog:

   (a) a sound file (so you can listen to the dialog)

   (b) a text transcript

1.3 A note about the examples

The examples in this manual come from a variety of sources, including daily conversation, previously collected spoken dialogs and other manuals. To avoid endless footnotes, we label each example with its source numerically, as follows:

1. Example comes from Mann and Thompson [4].
2. Example comes from a TRAINS-93 dialog [1].
3. Example is similar to utterances found in TRAINS-93 dialogs, but does not come directly from a TRAINS-93 dialog.
4. Example is a paraphrase or simplification of an extract from a TRAINS-93 dialog.

Some utterances come from text or from spoken monolog. Others come from dialog. The dialog examples appear in three columns: the first column indicates the speaker, the second the utterance number, and the third gives the utterance itself:

Example 1[4]

A	1	How many tankers do I have in Elmira now?
B	2	You have two tankers

2 Segmentation

The dialogs you will be looking at have been segmented (broken down) into "turns" and "utterances". A **turn** is a stretch of continuous speech by one speaker. The term "utterance" is not so clearly defined. Typically, dialogs are broken into **utterances** using rules similar to the following [1]:

- Break when the speaker changes
- Break when there is a syntactically complete unit (such as a clause or phrase), a breath or pause, and a "prosodic boundary"
- Break when there is a syntactically complete unit, a prosodic boundary, and a silence shorter than a pause
- Break when any two of the conditions in the second rule above are met

Roughly speaking, a **prosodic boundary** is a place where the speaker signals, without words but by adjusting aspects of the sound, that he or she is coming to a breaking point. For instance, most people raise the pitch of their voice when they finish a question, and lower it when they finish a statement. People frequently also speak more softly at the end of an utterance, and more loudly at the start of one.

You can see the segmentation into turns and utterances in the dialog transcript.

We will use utterances as "minimal spans" (the smallest dialog "chunks"). You will be marking relations that hold between short sequences of spans; usually two in a row, but sometimes three or more. When you mark a relation, the spans involved are combined to form a new (non-minimal) span. You will then mark relations that hold between sequences of those spans. You will end up with something that looks like a family tree, with only a few spans at the top and the individual utterances down at the bottom.

2.1 Re-segmenting

Usually utterances work fine as minimal spans. However, sometimes you may find yourself disagreeing with the existing segmentation.

Sometimes you may think a relation holds between two parts of the same utterance. In example 2a, by saying "I know you like Mahler" the speaker attempts to motivate the hearer to attend the concert. Here, you might want to break the utterance into two spans, as shown in example 2b.

Example 2a

A	1	You should attend the concert because I know you like Mahler

Example 2b

A	1	You should attend the concert
	2	because I know you like Mahler

Sometimes you may think that two utterances should be one span. In example 3, utterance 2 completes utterance 1. You might want to join these two utterances to form one minimal span.

Example 3

A	1	Have you ever read the Terry Pratchett book, um, "Witches"
B	2	"Witches Abroad"?
	3	I loved that book

You may change the segmentation as you annotate, if you note in your comments file the utterance(s) you modify and the reason for the modification. In general, you should ensure that the segmentation satisfies the following definition for minimal spans:

A **minimal span** must be one of the following, with earlier possibilities taking precedence over later ones:

1. A phrase separated from the immediately prior phrase by a "cue word" such as "because" or "since" (relationships between spans like these are intra-sentential relationships)
2. A syntactically complete stretch of continuous speech
3. A stretch of continuous speech ended by a pause, a prosodic boundary or a change of speaker

A **cue word** (or cue phrase) is a word or phrase that may signal the presence of a particular relation.

2.2 Timing

Sometimes you will notice that both dialog participants talk at once. In the transcripts, overlapping speech is indicated with +'s. For the purposes of this annotation, we must impose an ordering on minimal spans. We will say that **one utterance succeeds another if it starts after the other**.

Reordering utterances If two utterances overlap at their beginnings, and you think that the one listed second relates more closely to the previous material than the one listed first, then you can reorder the utterances. In example 4, utterances 2 and 3 start the same time and overlap. Because of the repetition ("Bath next") we can see that utterance 3 clearly is more closely related to utterance 1 than to utterance 2. So we might reorder utterances 2 and 3.

Example 4 [3]

A	1	Let's go to Bath next
	2	+ Then let's go + to Corning
B	3	+ Bath next mm-hm +

The only case where you may reorder utterances is if they start simultaneously.

3 When to stop

As you continue annotating, you will find yourself marking relations between bigger and bigger spans. Ideally, you would keep annotating until you had two or three spans covering the entire dialog, and then you would mark one relation between those spans. In practice, however, it is very difficult (and often not very useful) to achieve complete coverage of the dialog.

How do you know when you can stop annotating? You may think there is nothing more to do, but if you leave it for awhile you may see things you didn't see before. However, there are several conditions under which you may conclude you have a complete annotation:

1. You end up with one span that covers the whole dialog.
2. You reach a point where the structure between the spans is exactly the structure of the task the dialog participants are solving. For instance, say two people make plans to route a sequence of buses. You might end up having marked a sequence of small plans (one for each bus).
3. You find that there are only a few spans, and they correspond to a pattern resulting from how the dialogs were collected. In the case of many task-oriented dialogs, that might be three spans, one in which the task is described, one in which it is solved, and one in which the solution is summarized.
4. You reach a point where you cannot mark any relation.

If any of these conditions are met, you may stop annotating. **You should mark in your comments file the specific reason that you stopped**. If you have more than one top-level span, check with your supervisor or another annotator to ensure that you have not stopped too soon.

4 Marking relations

This section describes in general the relations that you can mark between spans in a dialog. For details about a specific relation, see section 6.

We have divided the relations into four groups: presentational relations; subject-matter relations; adjacency pairs; and schemas. **Presentational relations** serve to increase a desire in one or both dialog participants, such as the desire to act, or the inclination to like something or respect some belief. **Subject-matter relations** relate (part of) the content of one span to another, so they are also sometimes called informational relations. While presentational and subject-matter relations are commonly found in text, **adjacency pairs** are typically only found in dialog. They are conventional patterns of cooperative behavior, such as answering a question or acknowledging a greeting. **Schemas** are primarily structural (e.g. making a list) or contain large constructs (e.g. describing a situation).

It is not necessary for both participants in a dialog to recognize the purpose behind spans forming parts of a presentational relation. (In fact, sometimes a person might want to hide this purpose;

e.g. when a salesman is trying to convince or motivate you to buy a car you are unsure of, and yet avoid seeming to manipulate you.) It is, however, necessary for the participants to recognize the relations in the other categories. If they don't recognize the presence of a relation in one of these other categories, the dialog will seem incoherent to them.

For each set of spans, you should mark a subject-matter relation. You should also try to mark a presentational relation for each set of spans, and you should consider marking an adjacency-pair. If you can't mark any other kind of relation, consider marking a schema. If the appropriate relation of one kind covers a slightly different set of spans than the appropriate relation of another kind, mark both. At the next level of annotation, you will have to decide which possible resulting spans to use.

Occasionally, you may think more than one presentational relation or (rarely) more than one subject-matter relation is appropriate for a particular set of spans. **If you think more than one relation in a give category is present, select one to mark, and write the others in your comments file.**

The sets of relations we give in this manual are designed to cover everything you might encounter in task-oriented dialogs. However, if you find a relation that is not listed here, you may consult with your supervisor or other annotators about adding it to the list.

For each of the relations that we list in section 6, we include the name of the relation, a short definition, a test for applicability and some examples. We may also give other helpful information. Occasionally, we include some cue words or cue phrases that can signal the presence of the named relation [5]. However, it is very likely that the relation may be used but not be marked by any cue. Also, some cues may be used to signal any of several relations. **Do not use cue words/phrases as the sole indicator of the presence of a relation.**

For each category of relation, we have given either a decision tree (written as a list of questions) or a list of the relations in that category. You should use the decision trees and lists to help you identify appropriate relations.

4.1 Nuclearity

Most, but not all, of these relations hold between two spans. One of the spans will usually be more important in the context of the dialog than the other. You can test which one it is by asking yourself:

- Which span could I delete and still have the other make sense in this dialog?
- Which span could I replace and not lose any information in this dialog?

For instance, in the following short dialog utterance 2 is justification for utterance 1. The dialog would still make sense (would still be coherent) if utterance 2 were removed, but not if utterance 1 were removed:

Example 5

A	1	I'm going to call the company
	2	because I still didn't get my bill
B	3	Okay

When one span is more crucial to the coherence of the dialog than another, we call it the **nucleus**. The less-crucial span is called the **satellite**.

It is not always the case that one span is more important than the other. In the following dialog the participants are listing a sequence of events. You could not remove any utterance without defeating the purpose of the dialog:

Example 6

A	1	First put the oil pan ready under the oil cap
	2	Then unscrew the oil cap
	3	Wait for the oil to drain out
	4	Screw the cap back on
B	5	And then put the fresh oil in from the top of the engine?
A	6	That's right

It is possible for there to be more than one satellite in the spans for a relation, but as long as there is at least one satellite there cannot be more than one nucleus. That is, there are two possible patterns: multiple nuclei and no satellites, or one nucleus and one or more satellites.

The annotation tool provides a way for you to mark an utterance as a nucleus. **When you are marking relations, you should mark the nucleus or nuclei.**

In marking higher-level relations (relations involving non-minimal spans), it is acceptable to consider only the nuclei of the child spans when choosing relations. In the following example, utterances 3 and 4 form an *assertion-acknowledge* adjacency pair, in which utterance 3 is the nucleus. The resulting span, with utterance 2, forms an *enablement* presentational relation and a *process:step* subject-matter relation (utterance 2 is the nucleus). You can see this clearly by considering only utterances 2 and 3 (the nucleus of the non-minimal child span). Again, the span formed of utterances 2-4, together with utterance 1, forms another *enablement*, a *solutionhood* subject-matter relation, and a *proposal-accept* adjacency pair. This can be seen most clearly by considering only utterances 1 and 2 (the nucleus of the non-minimal child span).

Example 7

A	1	It says to find the slope
B	2	So you'll want to take the derivative
	3	So you'll need first to write down the formula
A	4	Okay

4.2 Leaving out utterances

Sometimes an utterance has been abandoned or is unintelligible, or both participants will start speaking together and then both stop without contributing anything to the dialog, or a participant may speak to the experimenter rather than to his or her partner. For these or other reasons, it may be necessary to **discount** an utterance. This means that you skip over it when marking relations, as if it weren't there. The annotation tool provides a way for you to mark an utterance as discounted. **You should mark discounted utterances.** Once an utterance has been marked discounted, you can choose to leave it out of the lowest-level relation you mark around it, or you may include it in that span so that it is no longer visible at the top of the tree.

Try not to mark an utterance as discounted unless absolutely necessary.

4.3 Marking content

For certain relations, it is necessary to give an indication of the type of the content. In these cases, you should mark the content type right after marking the relation. The annotation tool provides a way for you to do this. **It is important to read the relation definition to find out if you need to mark the content type for a given relation**.

Even if marking the content type is not required, you may choose to mark it if it provides you with more information for further annotation.

5 The annotation process

The basic annotation process you will use is as follows:

1. Listen to the entire dialog. You should be able to read the transcript while you are listening. You may have access to the same materials the original dialog participants used, such as maps, task descriptions and other information. While you are listening to the dialog, you can observe how the participants shared information with each other and developed their plan.
2. Starting at the beginning of the dialog, identify the relations that hold between different parts of the dialog. The procedure you should use to find and mark a relation is as follows:
   (a) Use the decision tree for the presentational relations to identify a presentational relation to mark.
   (b) Check the definition of the relation to make sure it is applicable and to clarify how to mark it.
   (c) Mark a content type if necessary.
   (d) Now, use the decision tree for the subject-matter relations to identify a subject-matter relation to mark.
   (e) Check the definition of the relation to make sure it is applicable and to clarify how to mark it.
   (f) Mark a content type if necessary.
   (g) Now, check the list of adjacency-pair relations. If you find an appropriate one to mark, do so.
   (h) Mark a content type if necessary.
   (i) If you cannot identify a presentational relation, a subject-matter relation or an adjacency-pair, then check the list of schemas to see if there is an appropriate one to mark.
3. If you have a question or comment or need to note a confusion, write it down in a comments file, with the utterance number or some other appropriate label (if you are doing a lot of re-segmenting, quoting the text may be preferable to using the utterance number). You should definitely make comments in the following circumstances:
   (a) If you re-segment utterances, note which ones.
   (b) If you think more than one relation in a given category (e.g. subject-matter relations) seems relevant, note the others.
   (c) When you stop annotating, note why.
4. Continue identifying relations until you have covered the whole dialog and satisfied the constraints in section 3.

1. In this set of spans, is one participant attempting to affect the beliefs of the other, the attitudes of the other, or the ability of the other to perform some action?
   - Belief – go to question 2
   - Attitude – go to question 3
   - Ability – *enablement*
2. Is one participant trying to increase the other's belief in some fact, or enable the other to better understand some fact?
   - Belief – *evidence*
   - Understanding – *background*
3. Is one participant trying to convince the other that she or he has the right to make some claim?
   - Yes – *justify*
   - No – go to question 4
4. Is one participant trying to increase the other's desire to perform some action?
   - Yes – *motivation*
   - No – go to question 5
5. Are the situations presented in the nucleus (N) and satellite (S) in contrast?
   - Yes – *antithesis*
   - No – *concession*

Figure 1: Decision tree for presentational relations

6 Relation descriptions

This section gives details about the relations you will use. In the descriptions and in the decision trees, S is the satellite and N is the nucleus.

6.1 Presentational/Intentional relations

These relations identify different ways that one dialog participant might try to influence the other. She or he may say something to affect the other's beliefs or attitudes, or to facilitate the other's ability to perform an action or understand a statement.

Most of these relations are more likely to be seen within a turn (or within multiple turns by one speaker), but they may also cross speakers. For instance, one participant may suggest *motivation* for performing a (shared or individual) action:

Example 1 [3]

A	1	(N) Well we could go via Delta instead
B	2	(S) Because it is shorter that way?

Or, one participant may request *background* on a situation. In the following, utterances 2 and 3 form a *question-answer*; the resulting span is *background* for the contents of utterance 1:

Example 2

A 1 (N) I need to get my hair cut
B 2 (S1) Why?
A 3 (S2) Oh because I don't like this style at all

Therefore, although we may use the terms "hearer" and "speaker" in the following definitions, the "speaker" need not be the speaker of both the nucleus and the satellite(s), and the "hearer" can, in some circumstances (such as example 1), be both participants.

The decision tree for the presentational relations appears in figure 1. When marking these relations, remember to follow the tree, and at the end confirm your choice by checking the relation definition and the examples.

The definitions for these relations come from [4]. The discourse cues come from [4] and [2].

Antithesis

Conditions:

1. The situations of N and S are in contrast.
2. One of the differences between N and S means that N and S can't simultaneously be considered to be true, valid, good or preferable.
3. The speaker views the situation of N as true, valid, good or preferable.
4. Knowing S, and knowing the incompatibility between N and S, will increase the hearer's regard for the situation of N as true, valid, good or preferable.

Cues: "on the other hand", "however"

Example 1 [4]

A 1 (N) We should go to Dansville first
2 (S) If we go to Bath first the other train will be late

Example 2

A 1 (N) We could buy a house
B 2 (S) Yes an apartment is so much harder to sell

Background

Conditions:

1. The hearer won't understand N properly without S.

Note: Before marking a *background* relation, check to make sure it is not really some other relation. The *background* relation should not in general be marked where an action or belief is being discussed, but only where N presents a simple fact.

Example 1
(N) I can't play any sports.
(S) When I was a child, I broke my leg and it never healed properly.

Example 2
(N) It was a gorgeous evening.
(S) There was a lovely sunset and we had such a good meal.

In example 3, there is a *background* relation between utterance 1 and the span consisting of utterances 2 and 3:

Example 3

A	1	(N) I am scared of loud noises
B	2	(S1) Why?
A	3	(S2) I used to live in Kosovo and there was always fighting

Concession

Conditions:

1. The speaker thinks the situation of N is true, valid, good or preferable.
2. The speaker is not saying that S is not true.
3. The speaker acknowledges a potential or apparent incompatibility between S and N.
4. The speaker thinks that the situations of S and N are compatible.
5. If the hearer realizes that the situations of S and R are compatible, the hearer's positive view of the situation of N will be increased.

Cues: "although", "though", "however", "even so", "nonetheless"

Example 1
(S) Although use of Primelene can lead to troubling side effects in some individuals,
(N) for most users, it is a life-saver.

Example 2

A	1	(S) We'll be really tired if we don't return until Sunday
	2	(N) The extra time in Israel will be worth it though

Enablement

Conditions:

1. N presents an (individual or shared) action that needs to be performed.
2. If the hearer knows S, he or she will be better able to perform the action of N.

Cues: "how" questions

Example 1
(N) You'll need to buy stamps.
(S) You can get them from the machine in the corner using these tokens.

Example 2 [3]

A	1	(N) You'll have to go to Corning first
B	2	(S1) Where is Corning?
A	3	(S2) Oh it's South of Bath

Example 3 [3]

A	1	(N) We have to make orange juice first
B	2	(S1) How do we do that?
A	3	(S2) Take the oranges to Corning where the orange juice factory is

Evidence

Conditions:

1. The hearer believes S or will find it credible.
2. If the hearer knows S, it will increase the hearer's belief in N.

Contrast with: *justify.*

Example 1
(N) The SlimQuick diet program really works!
(S) In only three weeks, I lost nine pounds, all while eating whatever I wanted.

In this example, utterances 3-5 are evidence for utterance 2:

Example 2 [4]

A	1	There's no other quicker way right?
B	2	(N) Uh I don't think so
	3	(S1) So the one that takes the longest is to make the orange juice
	4	(S2) and there doesn't seem to be a quicker way to do that
	5	(S3) because we have to go get the oranges and back again

Justify

Conditions:

1. If the hearer knows S, the hearer will be more likely to believe that the speaker has the right to say N.

Contrast with: *evidence*.

Example 1 [1]
(N) The next music day is scheduled for July 21 (Saturday), noon-midnight.
(S) I'll post more details later, but this is a good time to reserve the place on your calendar.

Example 2 [3]
A 1 (S) I have the paper right here
2 (N) and it says that there are only three boxcars

Example 3 [2]
A 1 (N) You'll have to tell me what it is
2 (S) I don't have their labels

Motivation

Conditions:

1. N presents an (individual or shared) action that needs to be performed.
2. If the hearer knows S, she or he will be more likely to want to perform the action of N.

Hint: If one of the spans is an invitation, offer or command, and you can put "because" in front of the other, this relation probably holds.

Example 1
(N) You should really come to the concert tonight.
(S) It's a once-in-a-lifetime chance to hear this performer.

Example 2
(S) We don't want to be late.
(N) So we'll not go to the store first.

Example 3 [4]
A 1 (S) So we have to make orange juice first
2 (N) So we're going to send engine two from Elmira to Corning to get the tanker

1. Is the relationship between the spans a causal relationship between an action/event and a situation? Does it answer the "why" question?
    - Yes – go to tree 1 (figure 3)
    - No – go to question 2
2. Do the spans give information about different situations/actions/objects, or only one?
    - One – go to tree 2 (figure 4)
    - Different – go to tree 3 (figure 5)

Figure 2: Decision tree for subject-matter relations

1. Is the action or event planned/meant?
    - Yes – go to question 2
    - No – go to question 4
2. Is the situation true at the time it is discussed (a present situation), or is it a future situation?
    - Present – go to question 3
    - Future – *purpose*
3. Is the action/event more central than the situation causing it, or is the situation more central?
    - Action/event – *volitional cause*
    - Situation – *volitional result*
4. Is the action/event more central than the situation causing it, or is the situation more central?
    - Action/event – *non-volitional cause*
    - Situation – *non-volitional result*

Figure 3: Subject-matter relations, tree 1

1. Does the satellite (S) give additional information about the nucleus (N)?
    - Yes – go to question 4
    - No – go to question 2
2. Is S a joke about N?
    - Yes – *joke*
    - No – go to question 3
3. Is the length of N the same as the length of S?
    - Yes – *restatement*
    - No – *summary*
4. Does N present a situation that S evaluates/measures?
    - Yes – go to question 5
    - No – go to question 6
5. Is the measure the speaker's subjective opinion, or is it objective?
    - Subjective – *evaluation*
    - Objective – *interpretation*
6. Does S contradict N?
    - Yes – *correction*
    - No – go to question 7
7. If the hearer did not know S but did know N, would he or she think the opposite of S was true?
    - Yes – *counter-expectation*
    - No – go to question 8
8. Does S give a solution to a problem presented in N?
    - Yes – *solutionhood*
    - No – go to question 9
9. Does N present a situation for which S gives a temporal or spatial framework?
    - Yes – *circumstance*
    - No – go to question 10
10. S gives additional factual detail about N. Use one of the following relations:
    - *set:member*
    - *abstract:instance*
    - *whole:part*
    - *process:step*
    - *object:attribute*
    - *generalization:specific*

Figure 4: Subject-matter relations, tree 2

1. Are the items being compared?
    - Yes – go to question 2
    - No – go to question 3
2. Are the items being:
    - Compared with respect to their differences? – *contrast*
    - Compared with respect to their similarities? – *compare*
    - Presented as alternatives? – *alternatives*
3. Is there a sequence of situations?
    - Yes – *sequence*
    - No – go to question 4
4. Can the contents of the spans be simultaneously true?
    - Yes – go to question 5
    - No – go to question 6
5. Are there two situations, one of which has to be true for the other to be true?
    - Yes – *condition*
    - No – *counter-expectation*
6. Does the contents of one of the spans solve a problem presented in the other?
    - Yes – *solutionhood*
    - No – *otherwise*

Figure 5: Subject-matter relations, tree 3

6.2 Subject-matter/Informational relations

These relations identify relationships between items in the contents of what one or more dialog participants said.

Task-oriented dialogs frequently involve the participants in making shared plans to perform or coordinate shared or individual actions. It is to be expected that both participants will contribute to the common pool of information, tasks, goals and actions. Therefore, subject-matter relations frequently span turns and speakers.

The decision tree for these relations starts in figure 2. When marking these relations, remember to follow the tree, and at the end confirm your choice by checking the relation definition and the examples.

Most of the definitions for these relations come from [4]. The discourse cues come from [4] and [2].

Abstract:instance

Conditions:

1. S presents an instance of, or example of, the concept described in N.

Contrast with: *set:member.*

Cues: "for instance", "for example"

Example 1
(N) The pongid family of the order primate is closely related to mankind.
(S) A particularly interesting example of the pongid is the chimpanzee.

Alternatives

Conditions:

1. At least two spans, all nuclei.

2. Each of the spans describes an alternative to the other spans.

Cues: "either", "or", "alternatively"

Example 1 [3]

A	1	(N) So we can't do this unless we send the boxcar first
B	2	(N) Or unless we attach two boxcars to one engine

Circumstance

Conditions:

1. S presents a situation that is true.

2. S provides a framework (temporal or spatial) for interpreting N.

Content types: When you mark the *circumstance* relation, you must also specify the type of the circumstance. The allowed types are:

- time-location ("before", "after", or a specific time)
- time-duration (e.g. "from 3 until 5")
- co-occurrence (e.g. "at the same time as...")
- space-location (e.g. "at my house")
- space-duration (e.g. "from my house to your house")

Contrast with: *condition, object:attribute.*

Cues: hypotactic clauses, e.g. X when Y, X where Y

Example 1
(N) I was in a bad car accident once,
(S) when I was still an inexperienced driver.

Example 2 [2]

A	1	(N) Determine the maximum number of boxcars of oranges that you could get to Bath by seven a.m. tomorrow morning
	2	(S) It is now midnight

Compare

Conditions:

1. No more than two spans (both nuclei).

2. The situations in the spans are the same in one or more ways.

3. The situations in the spans are compared with respect to one or more of these similarities.

Contrast with: *contrast.*

Cues: "also", "as well"

In the following example, utterances 2 and 3 form a *compare*, and together they form a *generalization:specific* with utterance 1:

Example 1

A	1	The family resemblance in Jane and Lisa is clear
	2	(N) They both have that high forehead
B	3	(N) They have the same nose as well

Condition

Conditions:

1. S presents a situation that is not currently true, e.g. is hypothetical, or in the future.

2. The situation of N cannot arise without the situation of S being true.

Contrast with: *circumstance, otherwise.*

Cues: "if", "what if", "assuming that", "unless", "until"

Example 1
(S) When you see Lucy,
(N) Jane is not far behind.
They are always together.

Example 2 [2]

A	1	(S) If we delay engine three for an hour
	2	(N) then then we're fine

Example 3 [3]

A	1	(S) What if the train gets too full?
B	2	(N) We'll just send another one

Contrast

Conditions:

1. No more than two spans (both nuclei).
2. The situations in the spans are the same in many ways.
3. The situations in the spans are different in a few ways.
4. The situations in the spans are compared with respect to one or more of these differences.

Contrast with: *compare.*

Cues: "on the other hand", "however"

Example 1

A	1	(N) My cat is fuzzier than your cat
B	2	(N) Well, my cat is more playful

Correction

Conditions:

1. Two spans, both nuclei.
2. One of the spans corrects some information contained in or inferable from the other.

Example 1 [3]

A	1	(N) Send the boxcar to Corning.
B	2	(N) We already sent it to Bath, remember?

Counter-expectation

Conditions:

1. S elaborates on (provides additional information about) the contents of N.
2. The hearer would otherwise think the opposite of S was true.

Cues: "however", "but", "nonetheless", "even so", "contrary to"

Example 1
(N) Contrary to what one would expect of a brain injury,
(S) many who have had suffered injury to the corpus callosum continue to function normally.

Evaluation

Conditions:

1. S expresses the speaker's subjective attitude about the information in N.

Contrast with: *interpretation.*

Example 1 [2]
A 1 (N) How about we go back to Elmira and then fill up with the orange
2 (S) Oh no no no gracious gracious bad bad move

Generalization:specific

Conditions:

1. S is a situation, object or action related to the material presented in N, but is more specific than N.

Contrast with: *set:member, abstract:instance, whole:part, process:step.*

Cues: "specifically"

Example 1
(N) My work is in the field of computer science.
(S) My current research is in artificial intelligence.

Interpretation

Conditions:

1. S evaluates the information in N according to some objective measure or framework of ideas not stated in N.
2. This measure is not the speaker's attitude or personal feeling.

Contrast with: *evaluation.*

Example 1
(N) We have had 30 days of record high temperatures so far this year.
(S) It is very warm.

Joke

Conditions:

1. S is a joke related to the content of N.

Example 1
A 1 (N) Then the train goes to Bath
B 2 (S) Train to Bath to take a bath....

Non-volitional cause

Conditions:

1. N presents an action or situation that is not a volitional (purposeful/intended) action.
2. S presents a situation that is the cause (but not the purpose) of the situation in N.
3. Without S, the cause of the situation or action in N would not be known by both dialog participants.
4. N is more central to the discourse purpose than S.

Contrast with: *non-volitional result.*

Cues: "because", "since"

Example 1 [1]
(S) Because we can mine more than we need,
(N) we've had plenty of manganese and iron ore for export.

Example 2
(S) It was stormy, and the boat was tossing.
(N) She felt ill.

Non-volitional result

Non-volitional results are generally side-effects of actions.

Conditions:

1. S presents a situation that is not a purposeful action.
2. N presents a situation that caused the situation presented in S.
3. N is more central to the discourse purpose than S.

Contrast with: *non-volitional cause*.

Cues: "consequently", "as a result"

Example 1 [1]
(N) The blast destroyed the plant and most of the surrounding suburbs.
(S) Several thousand people were injured, and about 300 are still in hospital.

Object:attribute

Conditions:

1. S is a property of the situation of N, a property of the subject matter of N, or a property of something related to N.

Content types: When you mark the *object:attribute* relation, you must also specify the type of the content of the attribute. The allowed content markings include (but are not limited to):

- location
- owner
- structure/shape
- contents
- size/capacity
- number
- type/class

Contrast with: *circumstance*. Be careful! This is a particularly tricky distinction. If S indicates the time or place of a situation, action or event use *circumstance*. Otherwise, consider *object:attribute*.

Example 1
(N) I have two saucepans.
(S) Each holds 3 quarts.

Otherwise

Conditions:

1. The situation of N is not currently true.
2. The situation of S is not currently true.
3. If the situation of N happens, the situation of S will not or cannot happen.

Contrast with: *condition*.

Cues: "otherwise", "or", "if not"

Example 1 [3]

A 1 (N) We have to send two boxcars
2 (S) Or we won't get them all

Example 2

(N) You should get down to Jones' car dealership now.
(S) Otherwise, you'll miss the best deal this year.

Process:step

Conditions:

1. S presents a step in the process described in N.

Contrast with: *sequence*.

In the example below, utterances 2 and 3 form a *sequence*, which is related to utterance 1 by *process:step*. (Under an alternative annotation, utterances 2 and 3 are both satellites related to utterance 1 by *process:step*.)

Example 1 [4]

A 1 (N) You want to move oranges there
2 (S1) What you'll have to do is you'll have to pick out an engine
3 (S2) And schedule a train to do that.

Purpose

Conditions:

1. N presents an activity.
2. S presents a situation that is not true yet but will become true as a result of performing the activity of N.

Cues: "in order to", "to"

Example 1
(N) She climbed a tree
(S) to see if there were houses near.

In the following example, utterances 2 and 3 form an *assertion-acknowledge*; the resulting span is the satellite for utterance 1.

Example 2 [2]

A	1	(N) Take those to uh Corning
	2	(S1) To pick up the oranges
B	3	(S2) Okay

Restatement

Conditions:

1. S rephrases N.

Contrast with: *summary.*

Cues: "in other words", "that is", "or rather"

Example 1 [3]

A	1	(N) Then take the train to Elmira
B	2	(S) First train to Elmira

Sequence

Conditions:

1. Two or more spans, all nuclei.
2. The spans in the relation describe a succession of situations, actions or events.

Cues: "first", "then", "next", "second", "after that", "finally"

Example 1
(N) First you jack up the car.
(N) Take off the old tire,
(N) then put on the new one.

Set:member

Conditions:

1. S presents a member of the set described in N.

Contrast with: *abstract:instance, generalization:specific.*

Cues: "one of"

Example 1 [3]
A 1 (N) We have three boxcars altogether
2 (S) The one at Bath is the closest

Solutionhood

Conditions:

1. S presents a problem.

2. N presents a solution to this problem.

Note: According to the original definition of this relation, "problem" is defined to include questions, requests for information and other requests, descriptions of desires, goals, gaps in knowledge or needs, and conditions that carry negative values such as frustrations. Be sure to also mark an adjacency pair, such as *question-answer, proposal-modify*, or *proposal-accept*, if permissible; they bring out these distinctions.

Example 1
(S) Are you coughing?
(N) Sinex can cure your cold!

Example 2 [4]
A 1 (S) We'll need a tanker for the orange juice
B 2 (N) Okay, when we're at Corning we'll take it back with us

Summary

Conditions:

1. S restates N in a shorter form.

Contrast with: *restatement.*

Cues: "to sum up", "in short"

In this example, utterances 1-3 form a *make-plan* (or a *sequence*). The resulting span, together with utterance 4, forms a *summary* relation:

Example 1 [3]

A	1	(N) So the first engine takes two boxcars of oranges from Avon to Corning
	2	And the second one takes two tankers to Corning to pick up the OJ
	3	And then it goes to Bath
B	4	(S) Oranges from Avon to OJ in Bath

Volitional cause

Conditions:

1. N presents a volitional (purposeful/intended) action, or a situation that came about because of one.

2. S presents a situation that could have caused the performer of the action of N to do that action.

3. Without S, the reason for performing the action of N, or the fact that the action or situation of N was intended, might not be evident.

4. The action or situation of N is more closely related to the discourse purpose than the situation of S.

Contrast with: *volitional result*.

Cues: "because", "in order to"

Example 1 [1]
(S) Writing had become almost impossible
(N) so we had the typewriter serviced
and I may learn to type decently after all these years.

Volitional result

Conditions:

1. S presents a purposeful action or a situation that could have arisen from a purposeful action.

2. N presents a situation that could have caused the situation of S.

3. N is more central to the discourse purpose than S.

Contrast with: *volitional cause*.

Cues: "since"

Example 1
(N) We were so poor,
(S) we ate cabbage for three months straight.

Whole:part

Conditions:

1. S presents a part of the situation of N, a part of the subject matter of N, or a part of something related to N.

Contrast with: *process:step*.

Example 1 [1]
(N) The 25th Conference on Artificial Intelligence will be held June 25-29 at Dover Beach, MD.
(S) Topics will include robotics, NLP and computer vision.

1. Greeting-response
2. Summons-response
3. Question-response
   (a) Question-answer
   (b) Question-no-answer
4. Assertion-response
   (a) Assertion-acknowledge
   (b) Assertion-modify
   (c) Assertion-reject
   (d) Assertion-hold
5. Proposal-response
   (a) Proposal-accept
   (b) Proposal-modify
   (c) Proposal-reject

Figure 6: Adjacency pairs

6.3 Adjacency pairs

For most people, these are easier to recognize than the relations in the other categories. They almost always cross speakers; for instance, one speaker asks a question and the other answers. However, sometimes they will be found within a turn. For instance, a speaker may ask a rhetorical question, or may ask a question to him or herself, or may ask a question and then realize that he or she already knows the answer:

Example 1
I mean, who would deliberately eat dirt?
Not me!

Example 2
A 1 (to self) So how shall I do this?
2 First the bank then the store then go home

Example 3 [3]
A 1 Where is Corning?
2 Oh yeah yeah down there I see

Some relations have sub-types that are more specific than the top-level relation. **If there are sub-types for a relation, mark the appropriate sub-type if possible.**

In general, these relations involve one nucleus and one satellite. However, sometimes you will find that it seems that neither satellite nor nucleus stands on its own, as in the following two examples:

Example 1 [3]

A	1	How many boxcars are there?
B	2	Two

Example 2 [3]

A	1	We could send the engine via Avon first
B	2	No, if we do that we'll never get there on time

In these cases, you may mark both (or all) spans as nuclei.

A list of adjacency pairs appears in figure 6.

Assertion-response

Conditions:

1. N contains an assertion.
2. S contains a response to N.

Subtypes:

- Assertion-acknowledge
- Assertion-modify
- Assertion-reject
- Assertion-hold

Contrast with: *proposal-response.*

This is an *assertion-acknowledge*:

Example 1 [3]

A	1	(N) And then the engine goes to Avon
B	2	(S) That's right

The following is an example of an *assertion-hold*:

Example 2 [3]

A	1	(N) We then go on to Bath
B	2	(S) Um...

Greeting-response

Conditions:

1. N contains a greeting or other social formula.
2. S is a response to N.

Example 1
A 1 (N) Hi
B 2 (S) Hello

Example 2
A 1 (N) Thanks
B 2 (S) No problem

Proposal-response

Conditions:

1. N contains a proposal: a command, offer, suggestion, or request for assistance (other than a request for information).
2. S contains a response to the proposal.

Subtypes:

- Proposal-accept
- Proposal-modify
- Proposal-reject

Contrast with: *assertion-response, question-response.*

This is a *proposal-accept*:

Example 1 [3]
A 1 (N) We could send the engine via Avon first
B 2 (S) Okay

The following is a *proposal-modify*:

Example 2 [3]
A 1 (N) We could send the engine via Avon first
B 2 (S) Yes, if we let it pick up two boxcars on the way

This is a *proposal-reject*:

Example 3 [3]
A 1 (N) We could send the engine via Avon first
B 2 (S) No, if we do that we'll never get there on time

Question-response

Conditions:

1. Two spans, (most frequently) both nuclei.
2. One span contains a request for information.
3. The other span contains a response, either an answer, or an indication that the speaker doesn't know the answer, or an indication of lack of understanding.

Subtypes:

- Question-answer
- Question-no-answer

Note: Many utterances that look like questions are **not** in fact requests for information. Prompts ("Hello?" "You there?" "Right?") and commands phrased as questions ("Could you shut the door?") are two particularly misleading such groups of utterances.

Contrast with: *summons-response, proposal-response.*

Content types: When you mark the **question-response** relation, you must also specify the type of the content of the question. The allowed content markings include (but are not limited to):

- yes-no
- time-location ("When")
- space-location ("Where")
- agent ("Who")
- reason ("Why")
- object ("What")
- process ("How")
- measure (number, size, distance, etc.) ("How much", "How far", "How long", etc.)

The following is an example of a *question-answer*:

Example 1 [3]

A	1	(N) How many boxcars are there?
B	2	(S) Two

This is a *question-no-answer*:

Example 2 [3]

A	1	(N) How many boxcars are there?
B	2	(S) I don't know

In this case, because A interrupts, we don't know if B would have answered or not, so this is a *question-response*:

Example 3 [3]

A	1	(N) How many boxcars are there?
B	2	(S) Um + I +
A	3	+ I + mean sorry how many engines

Summons-response

Conditions:

1. N contains a summons or prompt.
2. S contains a response to N.

Contrast with: *question-response*.

Example 1

A	1	(N) Can you hear me?
B	2	(S) Not very well

Example 2

A	1	(N) Right?
B	2	(S) Yes

1. Describe-situation
2. List
3. Make-plan

Figure 7: Schemas

6.4 Schemas

The relations described here are either specific to a particular domain, specific to task-oriented dialog, or are structural rather than rhetorical.

A list of these relations appears in figure 7.

If you think you find a new relation for this category, check with a fellow annotator or supervisor.

Describe-situation

Conditions:

1. At least one span. All spans involved are nuclei.
2. The spans all describe different sub-parts of the same situation.

Contrast with: *object:attribute, process:step, whole:part.*

In this example, the last three spans form a *describe-situation*, and the resulting span and the first span form an *evaluation*:

Example 1
The scene of the accident was a mess.
(N) There were four cars in the middle of the road.
(N) Glass and oil covered all four lanes.
(N) Traffic was backed up for three miles.

List

Conditions:

1. Two or more spans, all nuclei.
2. The spans form a conjunction or list.

Contrast with: *sequence.*

In the following example, utterances 2-4 form a *list*. The resulting span is related to utterance 1 by *whole:part* and *question-answer*.

Example 1
A 1 How many questions will be on the exam?
B 2 (N) There will be two review questions on differentiation,
3 (N) and three review questions on integration.
4 (N) There will be five more on new topics.

Make-plan

Conditions:

1. At least one span. All spans involved are nuclei.
2. The contents of the spans, taken as a whole, form a plan.

Contrast with: *sequence*.

The following utterances form the start and end of a sub-dialog involving the construction of a plan to route trains. Each step in the sub-dialog consists of an *assertion-acknowledge*; the series of spans thus created form a *make-plan*.

Example 1 [3]

A	1	Send one engine to Corning.
B	2	Okay.
...		
A	9	Also send another engine to Bath at the beginning.
B	10	Okay.
...		
A	15	And finally, back to Corning.
B	16	Right. All done.

7 ArgumentationTool

The annotation tool you will use, ArgumentationTool, is a simple Java application. You should be able to run it on any platform, if you download Java first.

7.1 Getting and installing ArgumentationTool

ArgumentationTool is available under the Gnu public license from:
`http://www.cs.rochester.edu/research/cisd/resources/aad/`.
So that we can keep you informed about bug-fixes and updates, please email *stent@cs.rochester.edu* when you download the tool.

To install ArgumentationTool, download it, unzip it, and compile the files. Complete installation instructions are available in the tool distribution.

7.2 Running ArgumentationTool

To start ArgumentationTool, type `java ArgumentationTool` at the prompt or run it from within your Java environment. When the tool has started, you will see several menus. Use the leftmost menu to load and save files.

7.3 New dialogs

Before you start annotating a new dialog, make sure the dialog transcript is in the right format. It must have all the utterances listed in order, and each utterance must start with "utt<number>:". If this is the case, you can open the file by going to the leftmost menu and selecting "New". Use the file-chooser to select the transcript file.

(If the dialog transcript is not in the required format, your site can write a tool to convert files into an appropriate format. See section 8 for more information.)

7.4 Loading and saving files

To save your work, go to the leftmost menu and click on "Save". Select or type the name of the file you would like to save your work in.

To open a file containing a dialog you have partially or completely annotated and saved, go to the leftmost menu and click on "Load". Select the name of the file you would like to load.

We recommend frequent saving.

7.5 Making and deleting spans; re-segmenting

Use the buttons just below the menus to make and delete spans.

To make a new span, select one or more spans and then click on "Make span". The spans you select should form a sequence with no gaps. You can select one span by clicking on it with the mouse; you can select multiple spans in a row by clicking on the first span, holding down the shift key and clicking on the last span.

To re-segment an utterance, select the utterance and then click on "Resegment". You'll see two copies of the text for the utterance. Double-click on the first copy, and delete some of it. Double-click on the second copy, and delete some of that.

To delete a span, select it and then click on "Delete span".

To expand a span (so you can see its parts), click on the little symbol just to the left of the icon for the span.

7.6 Moving spans

If you need to move more spans into a span, select the one you want to move, hold down the control key, select the destination parent, and then click on "Move span".

You have a lot of freedom to move spans around. Be careful not to confuse the ordering of utterances!

7.7 Nuclei

To identify a span as a nucleus, select the span and then click "Nucleify". Click again to undo this process.

7.8 Discounting utterances

To mark a span discounted, select the span and then click "Discount". Click again to undo this process. The "Discount" button should be used sparingly.

7.9 Marking relations

To mark a relation, select the span you would like to mark. Then go to the menu for the relation-type you would like to mark, and select the appropriate relation. It will be added to the annotated dialog. To remove that relation, simply select the span again and select that relation from the menu again.

Sometimes selecting a relation will open a little dialog box containing content-types. Select (or type in) the appropriate content type and click on "Set".

You can choose to add a content type even if none is required. Use the "Content-types" menu.

7.10 Making comments

Make your comments in the comment space at the bottom of the ArgumentationTool window. For each comment, note the utterance it applies to (if you are doing a lot of re-segmenting, quoting the text may be preferable to using the utterance number). If you save your annotation in the file `X.tree.txt`, the comments will be saved in `X.tree.txt.comments`.

You may also use a separate text editor to make your comments.

8 Moving to new domains

This section contains a brief review of those areas of the manual and tool that may need revision for a new domain or a different type of discourse.

8.1 The manual

When we started this process, we looked at several existing annotation schemes, including [4, 6]. In our first attempt at annotating dialog, we used [4]. We discovered that annotation schemes for marking rhetorical structure may need considerable modification before being used for new domains or genres. Most of this modification is in the area of finding a set of relations that is descriptive in the selected domain and that is mutually exclusive. For a new type of discourse (e.g. moving from text to speech or monolog to dialog), there may be other issues, such as segmentation and coverage issues; the issue of joint vs. individual intentions; and the issue of whether relations can cross speakers or turns. The process of developing this annotation manual is outlined in [8], which describes some of these difficulties.

Determining what relations are needed to mark a particular type of discourse in a particular domain is a time-consuming process. You can start with a standard set of relations ([3, 4]). Have one or two people mark a few dialogs/texts in the chosen domain or a closely related domain. Note the relations that are used, and leave out those that aren't. If one relation seems to be being used for a variety of phenomena, find or make more specific ones. If several relations are used to mark one phenomenon, remove all but one from the relation set. An alternative method for identifying a set of relations for a particular domain or genre based on discourse cues is outlined in [2].

If you use decision trees, you may need to write your own or modify the ones found in this manual. You may also want to add to the examples, although we have tried to provide a variety.

The presentational/subject-matter distinction is useful for a variety of different types of discourse, but adjacency pairs may not be necessary in monolog situations. Schemas may or may not be needed. Although [4] did not suggest marking both a presentational and a subject-matter relation for most sets of spans, we have found this to be very informative.

8.2 The tool

The first version of ArgumentationTool has been developed for our dialogs only. Future versions may be more general. Here we highlight some aspects of the tool that may pose challenges to others trying to use it.

Input files

There are two types of files that ArgumentationTool can read. The first type of file contains a sequence of utterances. Each utterance should begin with "utt<number>:"; utterances may span multiple lines. There may be a header before the first utterance, but after the first utterance there should only be more utterances. The TRAINS-93 transcripts (available at `http://www.cs.rochester.edu/research/cisd/resources/trains.html`) are in this format. This type of file is read using the "New" menu item.

The second type of file stores information in an XML-like format that contains everything necessary to reconstruct the tree. Specifically, each node is written out to file as two lines. The first line begins with "<" and ends with ">". In between these two characters, it has a set of "attr=val" pairs. The possible attributes include:

- number – line or utterance number
- id – a unique numeric id for this node
- nucleus – a boolean (true/false) value stating whether this node is a nucleus

- discounted – a boolean (true/false) value stating whether this node is discounted
- parent – the numeric id for the parent of this node. Nodes that are not yet in any relation have parent 0.

The second line contains the transcription for the utterance, or the relation label(s). A file in this format contains the nodes in depth-first order. Files of this type are read using the "Load" menu item.

It should be possible to write a simple Perl script that takes most transcripts and puts them in either the first or the second format. To convert a file to the second format, simply take each sentence/clause/utterance and add a line before it containing the relevant attribute-value pairs. The first line should be the root of the tree, so the text will be blank, the parent will not be given, and the id will be 0.

Relations

The relation types are fixed; you're going to have four menus plus a content menu no matter what (unless you modify the source code). However, the relations themselves are read from the relevant text files (`adjacent.txt`, `scheme.txt`, `content.txt`, `contentrelations.txt`, `present.txt`, and `subject.txt`). To alter the relations the tool shows, simply edit these text files.

References

[1] P. Heeman and J. Allen. The TRAINS 93 dialogues. Technical Report TRAINS TN 94-2, Computer Science Dept., U. Rochester, March 1995.

[2] A. Knott. *A Data-Driven Methodology for Motivating a Set of Coherence Relations*. PhD thesis, Department of Artificial Intelligence, University of Edinburgh, 1996.

[3] E. Maier and E. Hovy. Organizing discourse structure relations using metafunctions. In H. Horacek and M. Zock, editors, *New Concepts in Natural Language Generation: Planning, Realization, and Systems*. Pinter, London, 1993.

[4] W. Mann and S. Thompson. Rhetorical structure theory: a theory of text organisation. In L. Polanyi, editor, *The Structure of Discourse*. Ablex, Norwood, NJ, 1987.

[5] D. Marcu. *The Rhetorical Parsing, Summarization, and Generation of Natural Language Texts*. PhD thesis, Department of Computer Science, University of Toronto, 1997.

[6] M. Moser, J. Moore, and E. Glendening. Instructions for coding explanations: Identifying segments, relations and minimal units. Technical Report 96–17, University of Pittsburgh, Department of Computer Science, 1996.

[7] A. Stent. The Monroe corpus. Technical Report TR728/TN99-2, Computer Science Dept., University of Rochester, March 2000.

[8] A. Stent. Rhetorical structure in dialog. In *Proceedings of the 2nd International Natural Language Generation Conference (INLG'2000)*, June 2000. Student paper.

[9] D. Traum. Rhetorical relations, action and intentionality in conversation. In *Proceedings ACL SIG Workshop on Intentionality and Structure in Discourse Relations*, pages 132–135, June 1993.

[10] D. Traum and E. Hinkelman. Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, 8(3):575–599, 1992.

A Glossary

Adjacency pair A pair of turns or utterances, usually spoken by different speakers; the first turn or utterance imposes constraints on what can be said in the second turn or utterance.

Cue word/cue phrase A word or phrase that may signal the presence of a relation.

Dialog A spoken (or written) conversation between two participants.

Minimal span One of the following (earlier possibilities should be considered before later ones):

1. A phrase separated from the immediately prior phrase by a cue word such as "because" or "since" (relationships between spans like these are intra-sentential relationships)
2. A syntactically complete stretch of continuous speech
3. A stretch of continuous speech ended by a pause, a prosodic boundary or a change of speaker

Presentational relation A relation "whose intended effect is to increase some inclination" in the hearer, such as the desire or ability to perform some action or the belief in or acceptance of some situation or fact [4]. A presentational relation need not be recognized by both participants in a dialog.

Prosodic boundary Roughly speaking, a place where the speaker signals, without words but by adjusting aspects of the sound, that he or she is coming to a breaking point.

Relation Any relationship holding between two or more adjacent stretches of speech or text. These relationships may be primarily structural, conventional, intentional or informational.

Schema A structural relationship holding between elements of a dialog. Schemas are different from rhetorical relations and adjacency pairs, which express functional relationships between elements of a dialog.

Segment To break speech or text into meaningful units, such as paragraphs and sentences (text), or turns and utterances (speech).

Span A continuous stretch of speech or text that, either by structure, content or intent, forms a unit.

Subject-matter relation A relationship that holds between (part of) the contents of one span and (part of) the contents of another. Subject-matter relations differ from presentational relations in that they are only felicitous if they are recognized by both dialog participants.

Turn A continuous stretch of speech by one speaker.

Utterance A continuous stretch of speech by one speaker that forms (a fragment of) a syntactic phrase or clause or is bordered by a prosodic boundary or a pause.