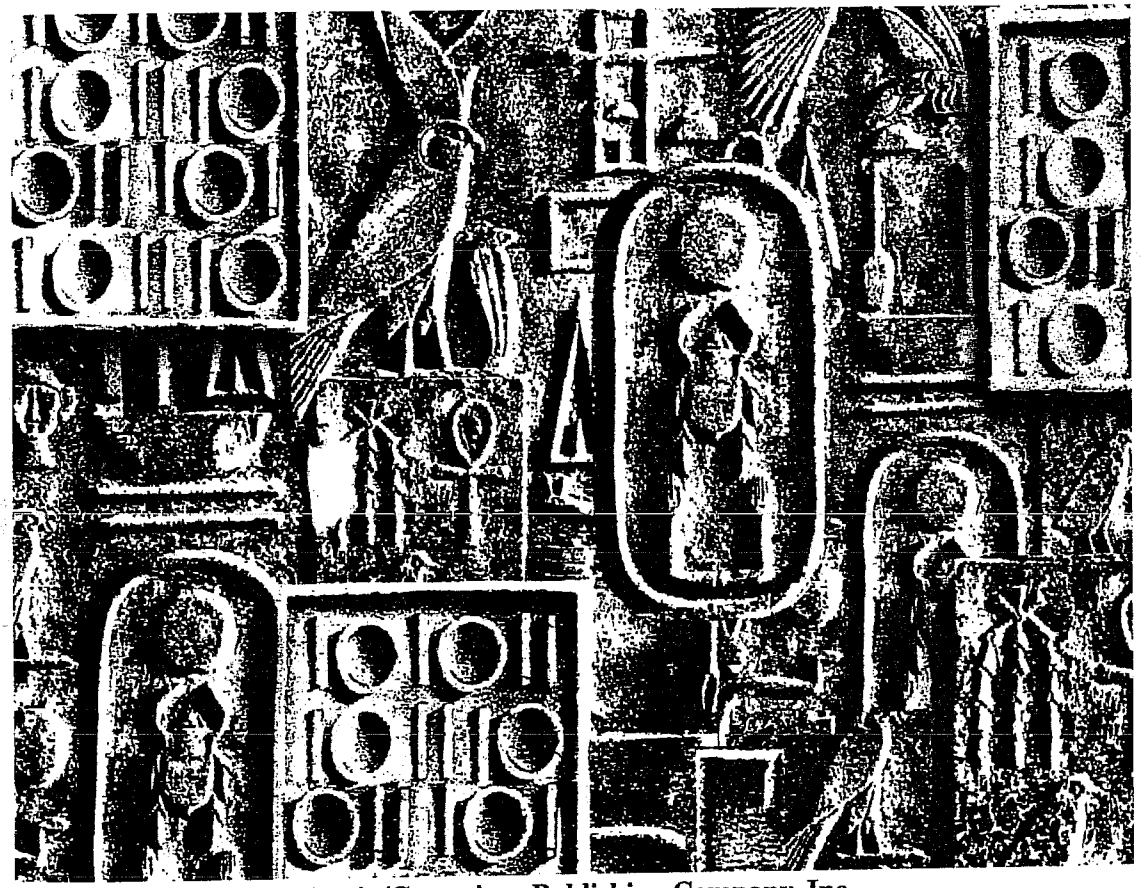
CHAPTERS 16.1 - 16.4
"DISCOURSE STRUCTURE"

Natural Language Understanding

James Allen
University of Rochester



The Benjamin/Cummings Publishing Company, Inc.
Redwood City, California • Menlo Park, California • Reading, Massachusetts
New York • Don Mills, Ontario • Wokingham, U.K. • Amsterdam • Bonn
Sydney • Singapore • Tokyo • Madrid • San Juan

Acquisitions Editor: J. Carter Shanklin
Executive Editor: Dan Joraanstad
Editorial Assistant: Melissa Standen
Cover Designer: Yvo Riezebos Design
Technical Assistant: Peg Meeker

Production Editor: Ray Kanarr
Copy Editor: Barbara Conway
Proofreader: Joe Ruddick
Design Consultant: Michael Rogondino

Macintosh is a trademark of Apple Computer, Inc. Word is a trademark of Microsoft, Inc. Canvas is a trademark of Deneba Software.

Camera-ready copy for this book was prepared on a Macintosh with Microsoft Word and Canvas.

Library of Congress Cataloging-in-Publication Data

Allen, James.
Natural language understanding / James Allen. -- 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-8053-0334-0
1. Programming languages (Electronic computers)--Semantics.
2. Language and logic. 3. Artificial intelligence. I. Title.
QA76.7.A44 1994
006.3'6--dc20

94-18218
CIP

7 8 9 10 – MA – 98

The Benjamin/Cummings Publishing Company, Inc.
390 Bridge Parkway
Redwood City, CA 94065

This chapter examines techniques for representing and reasoning about extended discourse beyond finding local connections between sentences as seen in Chapter 14. Extended discourse cannot be viewed simply as a linear sequence of sentences. Rather, in many cases the utterances cluster together into units, called segments, that have a hierarchical structure. Section 16.1 discusses the problems in viewing discourse as a linear sequence of sentences. Section 16.2 then defines the notions of discourse segments and cue phrases that signal segmental structure. Section 16.3 shows how the segment structure affects the interpretation of referential expressions, especially pronouns. Section 16.4 discusses how segments interact with inference to facilitate an understanding of the content of the discourse. Section 16.5 discusses tense and aspect, and shows how the interpretation of the temporal and causal connections between eventualities requires the use of segmental structure. Section 16.6 puts all the different components discussed earlier together to specify a model of discourse understanding. Section 16.7 presents an example that illustrates the issues discussed in the chapter.

16.1 The Need for Discourse Structure

The reference mechanisms presented in Chapter 14 were based on the structure of the previous sentence and on recency constraints. In dialogues where the topic may shift and change, however, you can see that these techniques are inadequate. Consider the following fragment, which could occur near the end of a dialogue between two persons, E and A, while E helps A assemble a lawn mower:

1a.	E:	So you have the engine assembly finished.
1b.		Now attach the rope to the top of the engine.
1c.		By the way, did you buy gasoline today?
1d.	A:	Yes. I got some when I bought the new lawn mower wheel.
1e.		I forgot to take my gas can with me, so I bought a new one.
1f.	E:	Did it cost much?
1g.	A:	No, and I could use another anyway to keep with the tractor.
1h.	E:	OK.
1i.		Have you got it attached yet?

The antecedent of *it* in sentence 1i is the rope last mentioned seven sentences earlier in 1b, even though objects mentioned since then, such as the gas can in sentence 1e, would satisfy any of the selectional restrictions that would be derived for *it*. Thus the history list mechanism fails to make the correct predictions in this dialogue. In fact, no simple generalization based on a linear ordering of discourse entities can provide a satisfactory solution. Intuitively, though, it is clear what is going on. Sentences 1c through 1g are a subdialogue incidental to the interaction involving attaching the rope. In 1h, E indicates that the current topic is completed by using the phrase *OK*. Thus in the interpretation of 1i, the relevant previous context is based on 1b. An account of this structure requires a notion of **discourse segments**—stretches of discourse in which the

sentences are addressing the same topic—and a generalization of the history list structure that takes the segments into account.

You might think that a generalization of the plan inference models derived in the last chapter might be useful for identifying the segments. Using such techniques, the system might be able to recognize that 1c is not a possible continuation of the plan to attach the rope, and thus represents a digression. Once the digression is completed, the plan recognizer could analyze 1i as querying the status of the action introduced in 1b. But trying to do all the work within the plan recognizer would be difficult. Whenever there is a shift of topic, such as at 1c and 1h, the plan reasoner would have to fail to find a connection between the old sentence and the new, and on the basis of this failure, initiate a new topic. This could be quite expensive, and might not be possible in some cases, since there might be an obscure interpretation that would allow a sentence such as 1c to be viewed as a continuation of the action described in 1b (for instance, the gasoline might be used to clean the engine before attaching the rope). Furthermore, you no doubt recognize that E explicitly told A that the topic had changed in 1c by using the phrase *By the way.* Such phrases, known as **cue phrases**, play an important role in signaling topic changes in discourse.

In addition, a plan-based model may not be appropriate in other conversational settings, such as debates, where intersentential relationships such as "sentence X supports the claim in sentence Y" or "sentence X contradicts the claim in sentence Y" may be relevant. Yet the same cue phrases could be used in this setting. These arguments lead to the conclusion that a theory of discourse structure cannot be explained solely in terms of action reasoning.

This chapter examines a model of discourse structure that allows each of the techniques discussed in the last two chapters to be generalized and integrated. The key idea is that a discourse can be broken down into discourse segments, each of which is a coherent unit and analyzable using techniques similar to those already presented.

16.2 Segmentation and Cue Phrases

While the need for segmentation of discourse is almost universally agreed upon, there is no consensus on what the segments of a particular discourse should be or how segmentation could be accomplished. One reason for this lack of consensus is that there is no precise definition of what a segment is beyond the intuition that certain sentences naturally group together. Notwithstanding these difficulties, a good model of segmentation is essential to understanding discourse. It divides the problem into two major subproblems: (1) What techniques are needed to analyze the sentences within a segment, and (2) how segments can be related to each other.

For practical purposes, a discourse segment consists of a sequence of clauses that display local coherence. The following properties should hold within a segment:

Event Described	Informational Relation	Communicative Goal
E1: Jack goes to store		Describe E1 as start of story
E2: Jack drives car	E2 part of E1	Elaborate on E1
E3: Jack buys lobsters	E2 before E3, E3 part of E1	Elaborate on E1
E4: Jack gets home	E4 provides temporal setting for E5	Elaborate story after E1
E5: Jack prepares for feast	E5 follows E4, E4 enables E5	Elaborate story after E4

Figure 16.1 Informational relations versus communicative goals

- Some technique based on recency (for example, a history list) should be usable for referential analysis and the handling of ellipsis.
- A fixed time and location characterize the clauses, or there is a simple progression of time and location (as in simple narratives).
- A fixed set of speakers and hearers are participating.
- A fixed set of background assumptions is relevant.

The last property requires that the modality of the text remains constant. For example, the text cannot switch from describing a sequence of actual events to describing a hypothetical event within a single segment.

These are the structural requirements on a segment. There are two approaches to characterizing what defines a segment. The **intentional** view is that all the sentences in a segment contribute to a common discourse purpose; that is, the same communicative goal motivates the speaker to say each sentence in the segment. The **informational** view is that all the sentences in a segment are related to each other by some temporal, causal, or rhetorical relations. For example, in narratives the sentences in a single segment should combine together to describe a coherent event or situation.

There is often a close correspondence between these two definitions. For instance, in most narratives, the writer's discourse intentions closely correspond to the informational level analysis. Consider the following start of a story:

2a. Jack shopped early in the day.
2b. He took his car
2c. and he bought a dozen live lobsters.
2d. When he got home,
2e. he spent the day preparing the feast.

Figure 16.1 shows an analysis at both the informational level and the intentional level. The informational relations tend to describe the "fine structure" of the

discourse, primarily how the events are causally and temporally related. The intentional level tends to address more global structural issues in terms of the communicative goal of relating the story.

While it seems that there is always an intentional level analysis, there are discourse situations where there is no informational analysis. In discourse 1, for example, sentence 1c does not have any informational relationship to the content of 1a and 1b. Rather, it is introducing a new topic for discussion. This can be analyzed at the intentional level as a change to a new topic. Examples like this motivate some researchers to argue that the intentional analysis is primary. But both levels are essential, and which approach seems more important will depend on the form of the discourse studied. In dialogues and debates, the intentional analysis seems most informative. In narratives and descriptive essays, on the other hand, the informational view seems most informative. Each view provides a useful analysis of the discourse, and neither can replace the other.

With this background, the issue of segmentation can be explored in more detail. We define a segment as a sequence of clauses, possibly interrupted by subsegments, forming a hierarchical structure. In some views, each clause forms its own primitive segment, and these segments combine to form larger segments. For reasons to be seen later, we will not take this view. A clause does not necessarily form a segment by itself, although it can do so under the right circumstances.

The most important aspect of segments is that they have a hierarchical structure. The explanation of why the pronoun *it* in sentence 1i cannot refer to the gas can mentioned in 1e depended on this fact. Since 1i is not in the segment defined by sentences 1c–1h, the gas can is not available as a discourse entity for 1i. Rather, sentences 1a, 1b, and 1i define a segment. The history list generated from these four sentences correctly predicts the antecedent for *it* in 1i.

This example shows an important function of discourse segments: They define the local context for the interpretation of referential expressions. The hierarchical structure then controls the availability of various different local contexts that might be used to process the current sentence.

The second important function of segments is to organize the information conveyed in a way that facilitates the identification of the relationship between a new sentence and the prior discourse. To support this identification process there must be some representation of the semantic content of each segment constructed by inferential processing.

Each segment is associated with a **local discourse state** (or simply discourse state), which consists of (at least) the following:

- the sentences that are in the segment
- the local discourse context, generated from the sentences in the segment, using techniques described in Chapter 14
- the semantic content of the sentences in the segment together with the semantic relationships that make the segment coherent

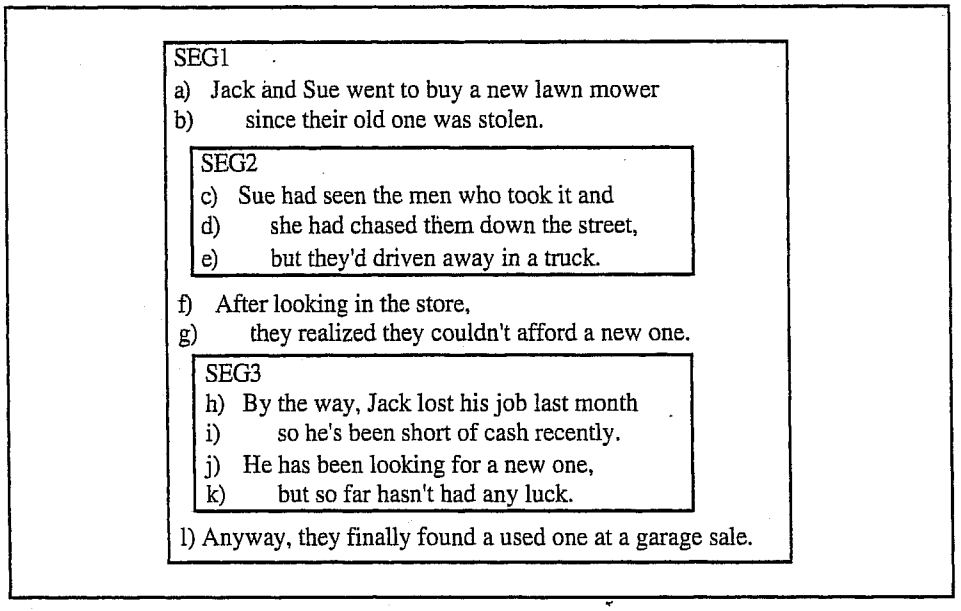


Figure 16.2 The segment hierarchy represented by boxing

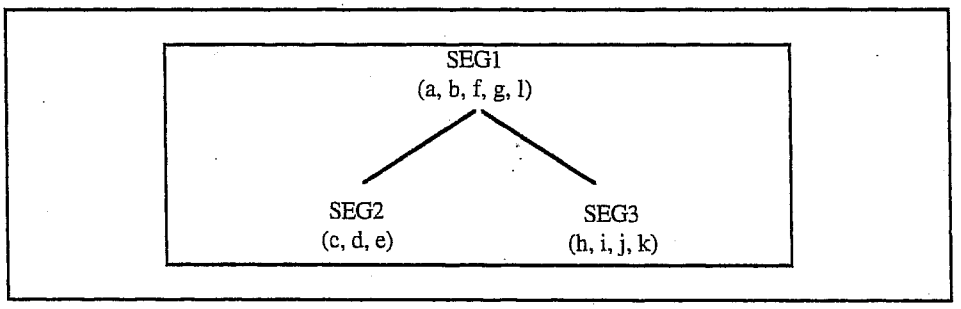


Figure 16.3 The same segment hierarchy represented as a tree

A complete discourse will typically involve many segments. Figure 16.2 shows the segmental structure of a dialogue represented by boxing of text. A segment is said to **contain** the segments that appear within it. The same information can be represented in tree form, as shown in Figure 16.3.

Additional concerns need to be addressed when considering on-line algorithms that understand discourse on a sentence-by-sentence basis. Such a processing model must be described in terms of extending an agent's representation of the discourse so far with a new sentence to create an updated representation of the discourse. This representation is called the **attentional stack** (or discourse stack) because it reflects what the agent is attending to in order to understand the next sentence.

The attentional stack consists of the discourse states reflecting the current structure of the ongoing discourse. The states on the discourse stack correspond

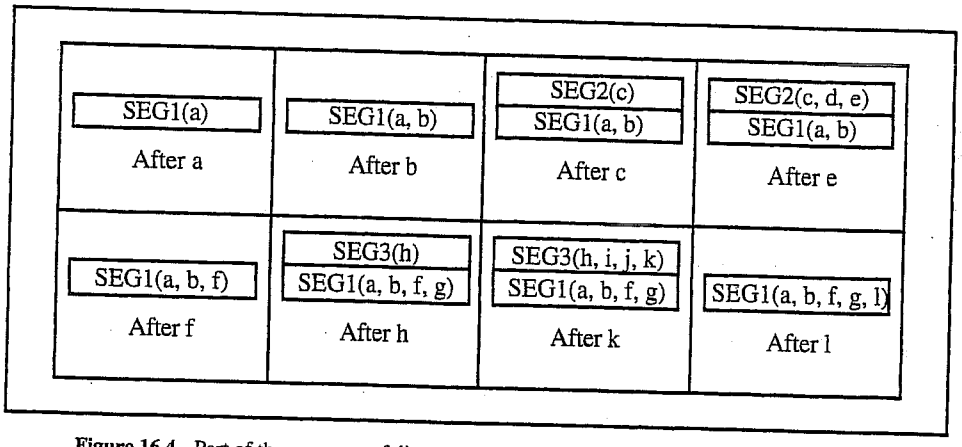


Figure 16.4 Part of the sequence of discourse stacks for the same discourse

to the set of segments that could be extended by the next clause. The top state on the stack corresponds to the most deeply embedded segment that can be extended. Each state on the stack corresponds to a segment that contains the segments of the states above it. To begin a new segment in the discourse, a new state must be pushed on the stack. To extend a segment corresponding to a state lower on the stack, all the states above it must be popped from the stack. This might sound complicated, but the stack model is quite simply related to the hierarchical segment structure. In particular, if you consider the sequence of discourse stacks, one after each clause, the sequence resembles a depth-first traversal through the tree of discourse segments. Figure 16.4 shows snapshots from the sequence of discourse stacks for the discourse shown in Figures 16.2 and 16.3. To show the relationship between the discourse states and their corresponding segments, the discourse states are labeled with the segment name followed by a list of the clauses seen so far in the segment. This provides a unique name for each of the discourse states produced during a discourse. For example, discourse state SEG1(a, b) is the discourse state corresponding to segment SEG1 at the point when clauses a and b have been processed. While not shown in the figure to save space, segment names will also need to include the completed subsegments so that the state of a segment can be tracked over multiple subsegments. Thus the full name for the state after clause f would be SEG1(a, b, seg2, f).

One of the more important indicators of the structure of a discourse is the use of cue phrases to signal the relationship of the next clause to the preceding discourse. Depending on the goals of their research, different researchers use different sets of cue phrases, but they can all be divided into two broad classes depending on what they signal. The first class identifies semantic relationships between clauses or states, and the second class indicates discourse structure directly without identifying a semantic relationship.

To introduce the first class, consider the two sentences

Jack went to the store. Sam stayed home.

Cue Phrases for Structure	Typical Use	Cue Phrases for Semantic Relations	Typical Use
anyway	end digression	and	continuation
by the way	start digression	because	causation/reason
bye	end dialogue	but	contrast
first	intro. subtopic (itemization)	furthermore	new subtopic
		however	contrast
incidentally	start digression	meanwhile	new topic (at same time)
last	new subtopic (itemization)	so	conclusion
next	new subtopic (itemization)	then	causal/ temporal
		therefore	summary
now	intro. subtopic	though	contrast
OK	close topic		

Figure 16.5 Some cue phrases and their uses

Taken as presented, there is no obvious relationship between the two sentences except an implied temporal overlap. But many words could be added to explicitly indicate the intended relationship between the events. For instance, you could indicate that the reason that Jack went to the store was because Sam stayed home,

Jack went to the store because Sam stayed home.

or that Sam stayed home because Jack went to the store,

Jack went to the store. So Sam stayed home.
Jack went to the store. Therefore, Sam stayed home.

or that Sam stayed home even though you would have expected him not to,

Jack went to the store but Sam stayed home.
Jack went to the store. However, Sam stayed home.

or that these two events are both evidence for some other conclusion,

Jack went to the store. Furthermore, Sam stayed home.

or finally, that a certain temporal relationship holds,

Jack went to the store. Meanwhile, Sam stayed home.

The second class of cue phrases signals the discourse structure directly without necessarily indicating a semantic relationship. Typically, they indicate segment boundaries. They include phrases used to end the current topic under discussion (such as *OK, fine*), to end the discourse itself (such as *bye, thanks*), to signal a digression (such as *by the way, incidentally*), to signal the end of a digression (such as *anyway*), or to indicate a particular discourse organization, such as itemization (for example, *first, second, next, last*). Figure 16.5 lists some cue phrases in these two broad classes together with some of their typical uses.

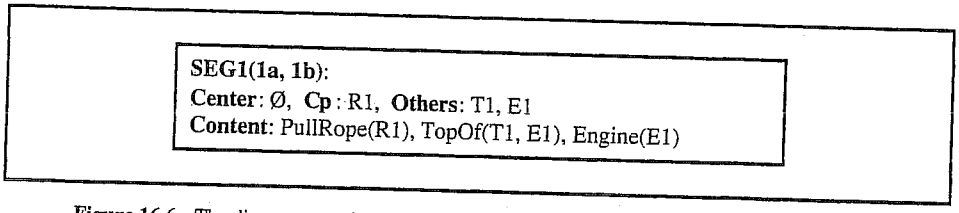


Figure 16.6 The discourse stack after sentence 1b

16.3 Discourse Structure and Reference

The example at the beginning of this chapter used a reference problem to motivate the need for a hierarchical discourse structure. The attentional stacks described in the last section provide the mechanism to account for this problem. The digression creates a new discourse state that temporarily hides the original discourse state when it is pushed on the stack. When the digression ends, its state is popped off the stack, and the original state becomes available. Consider this example in more detail. Figure 16.6 shows the discourse stack at the end of sentence 1b, *Now attach the rope to the top of the engine.* The current discourse entities and their properties are shown as part of the local discourse state for the segment. In particular, entities R1 (the rope), T1 (the top), and E1 (the engine) are available for subsequent reference, and R1 is the preferred next center. In sentence 1c, the cue phrase *By the way* signals a digression, so a new discourse state is pushed on the stack. This top state is then extended by utterances 1d through 1g, resulting in the stack shown in Figure 16.7. The discourse entities available for reference describe the new gas can, G3, and the tractor, T2. The clue word *OK* in 1h indicates the end of the digression, and the discourse state for SEG2 is popped off the stack, resulting in the discourse stack shown in Figure 16.8. Note that this state is the same as the one in Figure 16.6 that arose after sentence 1b. Thus when utterance 1i is processed, the pronoun *it* will refer to R1, as expected. It is not possible for it to refer to the gas can, G3, because that discourse entity is no longer available in the discourse state.

Pronouns play an important role in most arguments about discourse structure because, as you saw in Chapter 14, there are strong constraints on where the antecedent can appear. In particular, in most cases the antecedent is in the previous sentence and is subject to recency constraints. That is why examples such as the pronoun *it* in sentence 1i pose such a problem. The hierarchical discourse model provides an elegant solution to the problem that retains intuitions about the importance of recency.

The analysis for digressions is fairly straightforward because cue phrases typically signal the segment boundaries. Some other forms of discourse also have a clearly defined segment structure. For example, itemization constructs often use cue phrases to explicitly indicate their structure, as in the discourse

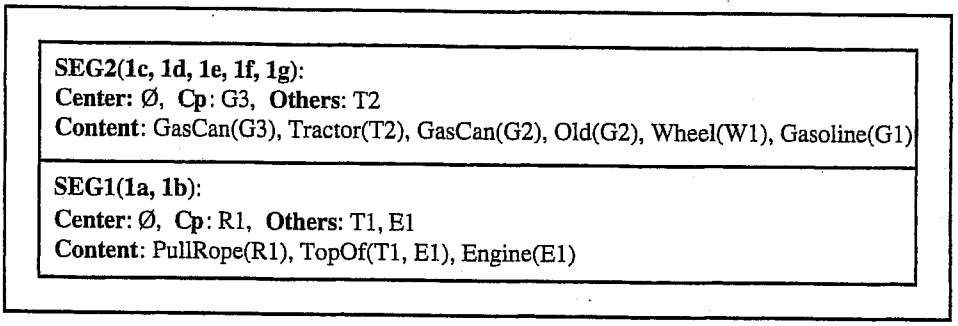


Figure 16.7 The discourse stack after sentence 1g

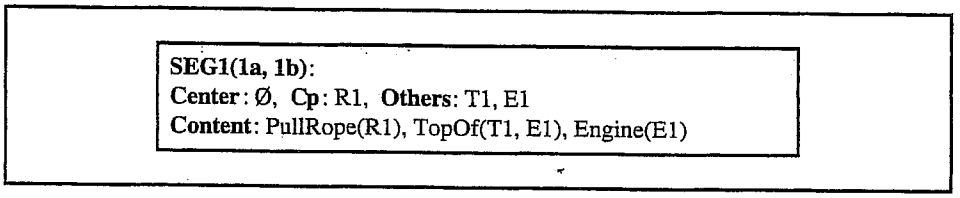


Figure 16.8 The discourse stack after sentence 1h

3a. There are many ways to identify a silver maple leaf.
3b. First, it has a silvery sheen on the back.
3c. If you hold it in your hand and move it,
3d. you will see the sun reflect off the back.
3e. Second, it has deep, pronounced notches between the points.
3f. The shape is quite similar to a red maple leaf.
3g. And third, if you break its stem, the sap will be milky.
3h. Break the stem and wait about 20 seconds and the sap should be visible.

This discourse has three subsegments. The beginning of each subsegment is explicitly marked by the cue phrases *first, second,* and *and third,* respectively. As each subsegment is popped and a new one begun, the discourse state for the segment consisting of 3a is used to identify the appropriate antecedent for the pronoun *it.* The discourse entities introduced in the previous subsegment are not available. For example, the pronoun *it* in 3g can't refer to the red maple leaf described in 3f, even though it is in the previous sentence. The segment structure of this discourse is shown in tree form in Figure 16.9.

Segmentation becomes more difficult to detect when a discourse has no explicit signal, but the topic slowly changes. It is sometimes difficult to distinguish between the case of a sentence that introduces a new segment and the case of a sentence that involves a natural progression of a topic occurring within a single segment. Such problems arise very frequently in narratives and stories, in which there is a continual progression of topic throughout. The story might describe a sequence of events over a long stretch of time and yet never show evidence of any hierarchical structure. If there is never an instance of popping

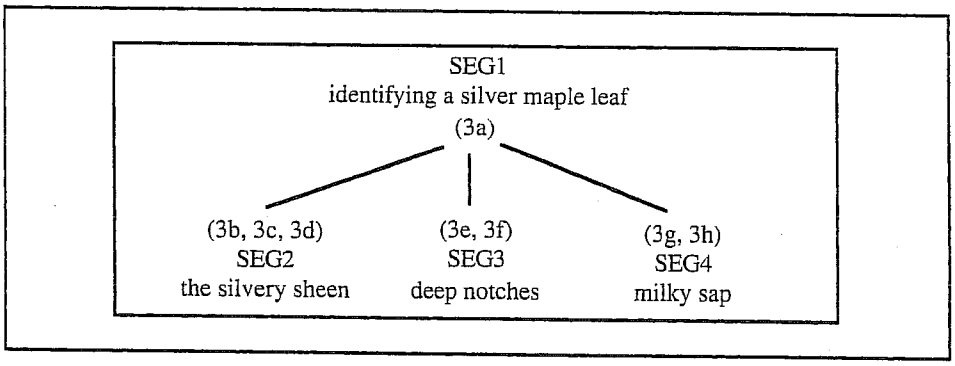


Figure 16.9 The structure of discourse 3

back to a previous segment, then you could argue that the entire story consists of a single segment.

The referential analysis itself can also suggest segmentation. In particular, if a referring expression is used that requires a discourse state lower on the stack in order to succeed, then this might force the completion of the segments corresponding to the states above it on the stack. Consider discourse 3 extended with the utterance

3i. These three tests are all you need to know.

Before processing 3i, the attentional stack contains the discourse states SEG1(3a, seg2, seg3) and SEG4(3g, 3h). Thus 3i could extend state SEG4(3g, 3h), or SEG4(3g, 3h) could be popped and 3i could extend SEG1(3a, seg2, seg3, seg4). There are no cue phrases in 3i to indicate a pop, but to successfully analyze the referring expression, SEG4(3g, 3h) must be popped. In particular, SEG1(3a, seg2, seg3, seg4) provides the discourse entities for the expression *These three tests,* which refers to the three methods described in SEG2, SEG3, and SEG4. The only context where all three tests could exist would be in the context SEG1(3a). Thus, in this case, it would be referential analysis that indicates a segment pop; based on the failure to produce a reasonable analysis of the referring expressions in the context SEG4.

This example also motivates another argument for segment structure. Consider what phrases evoke the discourse entities referred to in 3i by the NP *these three tests.* Interestingly, none are evoked by a single phrase. For instance, the first test is described by utterances 3b, 3c, and 3d. It is the combination of the content of all three utterances that evokes the discourse entity. This shows that segments themselves evoke discourse entities, presumably situations, that can be available for subsequent reference.

16.4 Relating Discourse Structure and Inference

As stated earlier, sentences within a segment must display local coherence. Depending on the form of the discourse, coherence may arise from causal

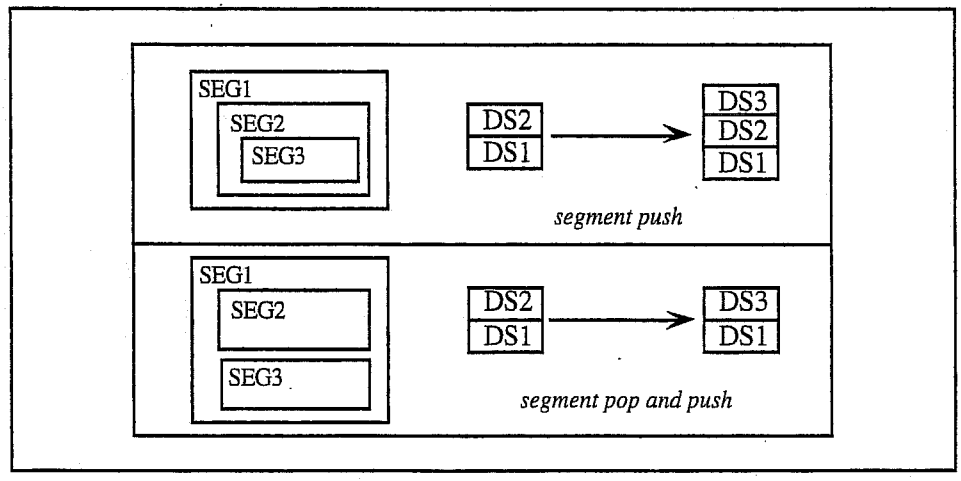


Figure 16.10 Two different ways a new segment can start

connections (as explored in detail in Chapter 15), or it may arise from other relationships such as evidence and counterevidence, as in a debate or argument. Whatever the mechanism for establishing local coherence, it is clear that this process will be strongly interrelated with the discourse structure. On the one hand, the segmentation of the discourse should be used to focus and direct the search for connections, and on the other hand, finding connections (or failing to find connections) can affect decisions about segmental structure. This section explores the relationship between discourse structure and inference.

You have seen several examples of segmentation so far. When considering inferential processing, these examples fall into two very different classes. Discourse 1 shows an example of a subsegment that is a digression. In this case there is no inferential connection between the segments. Each segment is treated independently and has its own coherent structure. Discourse 3 shows the opposite extreme. Here there would be a one-to-one correspondence between the discourse structure and a representation of the task of identifying silver maple leaves using the three tests.

Given this range of behavior, it might seem that there is little to say about how discourse structure and inferential structure relate. There is, however, an important constraint between the two that helps decide what segmentation structures are possible. Specifically, whenever a sentence begins a new segment, there is an ambiguity between whether the new segment ends the prior segment or is a subsegment of the prior segment. Figure 16.10 shows this choice in terms of its effect on both the segmental structure and the change in the discourse stack. The discourse state for each segment SEG_i is written as DS_i. The decision affects what discourse entities are available for reference and how the discourse could continue once the new segment ends.

The noncommittal approach would be to always push new segments, thereby allowing any previous topic to be resumed later. This approach certainly

seems to work for digressions and interruptions, but it goes against intuition for the structure of discourses such as 3. If SEG3 was a subsegment of SEG2 in discourse 3, the sun mentioned in 3d would still be available for later pronominal reference. In this discourse, the cue phrases explicitly signal the desired structure, so this interpretation is not possible.

Unfortunately, in many cases there are no explicit signals to indicate which interpretation is intended. In these cases the distinction is made based on inferential grounds. Specifically, if an inferential connection can be identified, we may know that a segment is completed because the discourse has moved on to the logical next topic. For instance, in a discourse that is describing some event, segment pushes might correspond to a decomposition relationship between the events, while a segment pop and push might correspond to moving to the next event in a sequence.

While all discourses clearly are not descriptions of events, the hierarchical structure is useful for many different forms of discourse. To unify these approaches across domains, the notion of the **discourse purpose** of a segment is introduced. While defining this notion precisely is difficult, the intuitions are clear. The idea is that the constraints between segments arise because of why the speaker is saying the sentences in each segment. The claim is that a segment push occurs when the new segment (corresponding to DS3) is said to accomplish a subgoal of the goal of the current segment (corresponding to DS2). A pop and push, on the other hand, occurs when the goal of the new segment is not a subgoal of the goal of the previous segment, but is a subgoal of the goal of the segment embedding the previous segment. The specification of what goals are suitable as discourse purposes is defined by the type of discourse.

For instance, in a discourse in which the conversational goal is to describe a (possibly complex) event, the discourse purpose hierarchy might correspond to the event decomposition hierarchy. In particular, the discourse purposes would all be of the form "describe event X." If event X is part of event Y, the discourse purpose "describe event X" would be a subgoal of the discourse purpose "describe event Y."

Consider another form of discourse, namely debates. The discourse purpose is to establish some claim and the subgoal relationship corresponds to evidential support. For example, if you have a goal to establish claim X, then a subgoal might be to establish claim Y, where Y would tend to make the other person believe X, that is, where Y provides evidence for X. With this correspondence a segment push may correspond to the case where the new segment has a discourse purpose of establishing a claim made in the prior segment.

As one final example, consider a discourse whose purpose is to describe all the rooms in a house. The discourse purpose hierarchy could reflect the physical layout of the house, for example, describing the living room is a subgoal of describing the house, and describing the alcove in the living room is a subgoal of describing the living room.

With this abstraction away from the specific inference process underlying the discourse, we can now state the constraint between discourse structure and inference in general terms. We say that a discourse purpose DP1 **dominates** another discourse purpose DP2 if and only if DP2 is viewed as a subgoal of DP1. A discourse purpose DP1 **immediately dominates** a discourse purpose DP2 if and only if DP1 dominates DP2 and there is no discourse purpose DP3 such that DP2 is a subgoal of DP3 and DP3 is a subgoal of DP1.

It is important to remember that not all segments must be related to each other in these ways. In fact, segments may be unrelated to each other, as seen in interruptions and digressions. So the fact that a segment is contained within another segment does not mean that the discourse purpose of the first must dominate the discourse purpose of the second. Rather, the constraints work in the opposite direction. When the inferential component identifies a dominance relationship, this imposes a constraint on the possible attentional stacks (and hence segmental structure). In particular, the following constraint holds:

> **Domination Constraint**—If the discourse purpose of the segment associated with discourse state DS1 immediately dominates the discourse purpose of the segment associated with DS2, then if DS2 is on the attentional stack, DS1 must be immediately below it on the stack.

This constraint forces attentional stack updates that agree with intuition. For example, consider the following discourse, which is a revised version of discourse 3 with the cue phrases removed:

4a. There are many ways to identify a silver maple leaf.
4b. It has a silvery sheen on the back.
4c. If you hold it in your hand and move it,
4d. you will see the sun reflect off the back.
4e. It also has deep, pronounced notches between the points.
4f. It is quite similar to a red maple leaf.
4g. If you break its stem, the sap will be milky.
4h. Break the stem and wait about 20 seconds and the sap should be visible.

Removing the cue phrases makes the discourse a bit more difficult to understand, but it is still comprehensible. The desired segment structure is still as shown in Figure 16.9; that is, the top level segment for the discourse is SEG1, and there are three subsegments, SEG2, consisting of 4b, 4c, and 4d, SEG3 with 4e and 4f, and SEG4 with 4g and 4h. In this setting the discourse purpose is to convey how to perform the task of identifying silver maple leaves, and the domination relationship is defined by the task/subtask relationship. Assuming the appropriate domain reasoner, the discourse purpose of SEG1 immediately dominates the discourse purposes of all three of SEG2, SEG3, and SEG4. Thus the domination constraint uniquely determines the progression of discourse stacks. Figure 16.11 shows the discourse stack update given sentence 4b. Given the domination constraint, there

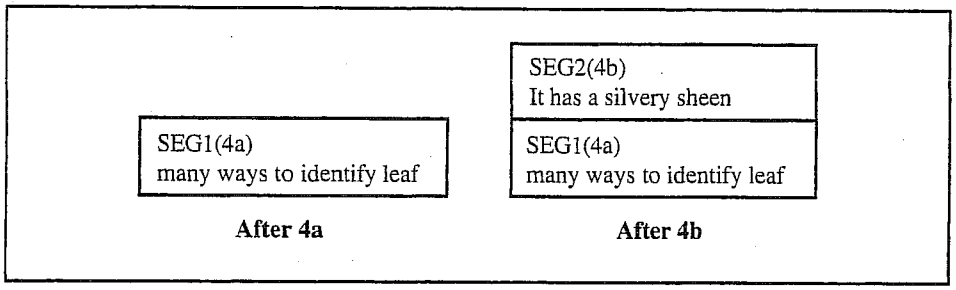


Figure 16.11 The attentional stack update given sentence 4b

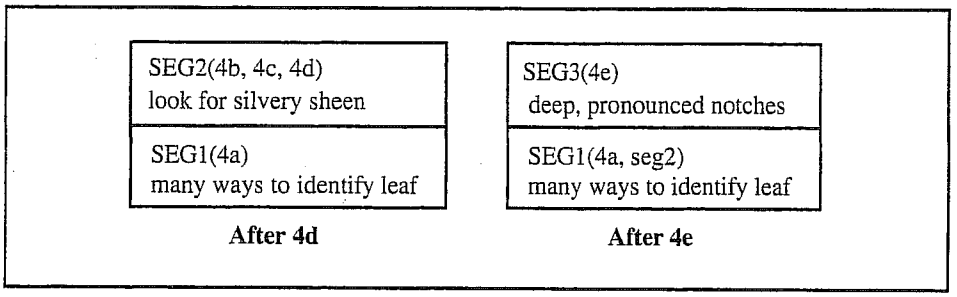


Figure 16.12 The attentional stack update given sentence 4e

is no other possibility. In particular, the discourse state SEG1(4a) could not have been popped off the stack because its purpose dominates the purpose of the new segment started by 4b.

Figure 16.12 shows the attentional stack update for (4e). Again, if the domain reasoner identifies that SEG1 immediately dominates the new segment SEG3, then exactly one transition satisfies the domination constraint. The state SEG2(4b, 4c, 4e) must be popped off the stack so that SEG3(4e) can be immediately above SEG1(4a, seg2).

Thus, to use this model in a particular application, you must determine what inferential connections in your domain induce the immediate dominance relationship between discourse segment purposes. Once this is determined, the domination constraint will allow you to use the inferential processing to restrict the possible discourse structure and attentional stack updates.

Another important advantage to the discourse purpose approach is that it allows the analysis of discourses that may be defective in some way. For instance, it might be that silver maple leaves do not have a silvery sheen on the back, and the speaker is mistaken. However, the analysis of the discourse remains the same because the discourse purposes are the same, even if they are based on erroneous beliefs. Of course, to recognize the structure of such dialogues, you must be able to identify the intended purpose even when it is based on an erroneous belief. This is a very difficult problem in general, but the theory leaves the door open for future research. Note that in discourse 3, with the structure made explicit by cue phrases, the constraints could be used in reverse. A system could

use the domination constraint to infer from the discourse structure that the speaker believes that silver maple leaves have a silvery sheen on the back, since this is the most plausible explanation for the way the discourse is structured.

16.5 Discourse Structure, Tense, and Aspect

Tense and aspect provide a rich source of information relating events within a discourse segment as well as providing constraints that can be used to identify segment boundaries. This section considers how tense and aspect affect both segmentation and the inferential processes used to derive the connections required to establish coherence.

Consider first the effect of tense and aspect within a single segment. In simple narratives within a single segment, if a sequence of sentences describes a series of events, then the events occurred in the same order as they are described. This is sometimes called the **narrative convention**. Consider the discourse

5a. Jack went to the store.
5b. He bought some roses.

The intuitive reading is that Jack bought the roses after going to the store. Some researchers have suggested using this constraint to improve their inference algorithms by only considering expectations that involve events after the last described event. Unfortunately, this hope is unfounded except in the simplest domains. Consider the discourse

6a. Jack went to the store.
6b. He walked there along the river.

In the most natural interpretation, the event of walking along the river occurs as part of going to the store. Or consider the discourse

7a. Jack showed us his new car.
7b. He bought it at Honest John's Auto Mart.

In this case the event described in 7b clearly precedes the event described in 7a.

It might appear that any relationship can hold between the events in two simple past tense sentences, but this is not the case. The relationships seem to be limited to a few specific causal relationships plus a default temporal reading. We capture this fact with a new relation called the **orients relation**. Any two consecutive eventualities in the same segment will have an orients relation between them. A reasonable first attempt at defining the orients relation is as follows:

If an event E_1 orients an event E_2, then
1. If E_2 is part of E_1, then $E_2 \subseteq E_1$ (i.e., E_2 occurs during E_1).
2. If E_2 enables E_1, then $E_2 <: E_1$ (i.e., E_2 precedes E_1).
3. Otherwise, $E_1 <: E_2$ (i.e., E_1 precedes E_2).

Condition 3 gives the default case, encoding the narrative convention.