On the Coherence and Structure of Discourse

Jerry R. Hobbs

1 Discourse is Coherent

Let us begin with a fact: discourse has structure. Whenever we read something closely, with even a bit of sensitivity, text structure leaps off the page at us. We begin to see elaborations, explanations, parallelisms, contrasts, temporal sequencing, and so on. These relations bind contiguous segments of text into a global structure for the text as a whole.

Consider a specimen:

(1a)   I would like now to consider the so-called "innateness hypothesis,"
(1b)   to identify some elements in it that are or should be controversial,
(1c)   and to sketch some of the problems that arise as we try to resolve the controversy.
(2)   Then, we may try to see what can be said about the nature and exercise of the linguistic competence that has been acquired, along with some related matters.

Chomsky, *Reflections on Language*, p. 13.

Between sentence (1) and sentence (2) there is a temporal relation, indicated by "then", linking two topics Chomsky intends to discuss. Clause (1a) states the first topic, and clauses (1b) and (1c) elaborate on that by breaking it into two subtopics that will be discussed in sequence. This structure may be represented as in Figure 1. One could of course argue about details of this analysis; in fact, one of my aims in this paper is to develop a *way* of arguing about the details.

Numerous other researchers have pointed out that such relations exist. Robert Longacre has a chapter in *Anatomy of Speech Notions* (1976) on "combinations of predications", among which he includes conjunction, contrast, comparison, alternation, temporal overlap and succession, implication,
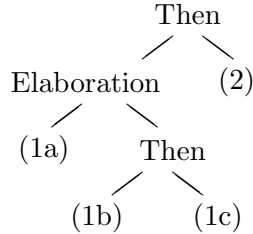
1

Figure 1: Structure of sentences (1)-(2).

and causation. Joseph Grimes has a chapter in *Thread of Discourse* (1975) on these relations; his list includes alternation, specification, equivalence, attribution, and explanation. Others have proposed similar lists. Grimes calls these relations "rhetorical predicates", as do Mann and Thompson in their recent work (1986). Fillmore (1974) has called them "sequiturity relations". Edward Crothers (1979) calls them "logical-semantic connectives". In accord with the tradition of using idiosyncratic terminology, I will call them "coherence relations". Mann and Thompson (1986) and Hovy (1988) have begun to give them more formal definitions.

The question is, what are we to make of these relations? Most authors have only pointed out their existence and listed, largely without justification, the relations most often found in texts. Longacre and Grimes describe the relations carefully. Crothers attempts to correlate the types of texts with the frequency of the relations that occur in them. Bonnie Meyer (1975), building on Grimes' work, classifies texts according to the structure their coherence relations impose and tries to relate that to what people remember of passages.

In this paper I attempt to embed a theory of coherence relations within the larger context of a knowledge-based theory of discourse interpretation. The process of interpreting discourse is a process of using our knowledge gained in the past to construct a theory of what is happening in the present. We therefore need, first of all, a framework in which we can talk precisely about the representation and use of knowledge. Such a framework is sketched in Part 2. An important aspect of one's understanding of a discourse is recognition of the coherence relations in it.[1] Part 3 is an ac-

---

[1]Here and throughout I intend "recognition" to refer not to conscious recognition, but to the implicit or latent sort of recognition that occurs, for example, when one "recognizes" the syntactic structure of a sentence. A similar remark applies to "inference".

count of the coherence relations, in which their intimate connection with the knowledge of the speaker and listener is explored. Of particular concern is the problem of giving formal definitions to the coherence relations in terms of inferences drawn by the listener. In Part 4 it is shown how larger-scale structures in discourse are composed out of the coherence relations. This will help elucidate the elusive notions of "topic" and "genre", and allow us to examine some of the ways in which ordinary discourse is often incoherent. Thus, in Part 3 we examine the internal structure of the coherence relations and in Part 4 the structure they impose on the text as a whole. In Part 5 a method for analyzing discourse is suggested, which allows the structure of discourse and its underlying knowledge to illuminate each other.

2 The Framework for a Theory of Discourse Interpretation

We understand discourse so well because we know a lot. Therefore, a theory of discourse interpretation must first and foremost be a theory of how knowledge is used in solving the interpretation problems posed by the discourse. This and other considerations suggest that the very large problem of discourse interpretation be carved into the six (still very large) pieces, or subtheories, listed below. Each subtheory is illustrated with an example relevant to one interpretation problem—the resolution of the definite noun phrase "the index" in the following text:

(3a) John took a book from the shelf.
(3b) He turned to the index.

  1. **Logical Notation,** or **Knowledge Representation:** We must have a logical notation in which knowledge can be expressed and into which English texts can be translated. This problem has of course given rise to a large area of research, but I think the difficulties have been overstated. Typically, workers in this field have been trying not only to represent knowledge, but to do so in a way that satisfies certain stringent ontological scruples and canons of mathematical elegance, that lends itself in obvious ways to efficient computer implementation, and explains a number of recalcitrant syntactic facts as a by-product. If we decide to ignore these criteria or let some other part of the total system bear their weight, then most (though not all) of the problems of knowledge representation evaporate.

For our example, let us suppose our logical notation is simply first-order predicate calculus. It allows us to make and combine predications and provides us with a rich set of predicates, such as *book* and *index*.

2. **Syntax and Semantic Translation:** Texts must be translated, sentence by sentence, into the logical notation. This also has been a major area of research for over a decade in linguistics and computational linguistics (Montague, 1974; Woods, 1970), and the solution has largely been worked out. The processes to be used are clear, the most commonly encountered syntactic constructions have been adequately analyzed, and current research is for the most part concerned with second-order refinements.

In our example, we may assume that syntax and semantic translation produce a logical form for sentence (3a) that includes the expression

$$book(b)$$

and for sentence (3b) a logical form that includes

$$index(i, z),$$

where b, i, and z are existentially quantified variables. The processes of syntax and semantic translation can not be expected to determine what z is, that is, that i is the index of b. That happens below.

3. **Knowledge Encoding:** The knowledge of the world and the language that is required to understand texts must be encoded in what may be called a "knowledge base". It will necessarily be huge, and the project of determining what needs to be represented, how to encode and organize it, and whether or to what extent it is consistent is correspondingly huge. Whether or not the project is tractable remains to be seen, but it is currently a healthy area of research (see Hobbs and Moore, eds., 1984; Hobbs et al., 1987; Lenat et al., 1986; Dahlgren, 1985). We need not wait for the completion of this research before proceeding to a theory of discourse interpretation, for we can make general assumptions about how the knowledge is encoded, and we can assume specific (but not too specific) facts to be present in the knowledge base, as convenient. This is a way of isolating the problem of interest—how those facts are *used* by the interpretation processes.

In our example, we may assume that one of the facts in the knowledge base is the fact that books (at least sometimes) have indexes:

$$(\forall x)(\exists y)\, book(x) \to index(y, x)$$

4. **Deductive Mechanism:** If we are to use knowledge stored as "axioms" in a logical notation, we must have some sort of theorem prover, or

deductive mechanism, to manipulate these axioms and to draw appropriate conclusions. This is not to say that language understanding *is* deduction, rather that it *uses* deduction. Deduction must be under the strict control of the processes described below.

Automatic theorem proving is also a healthy area of research. There are many who have despaired of the possibility of devising efficient deductive procedures, but I think that despair is premature, for two reasons. First, parallel machine architecture, which the human brain surely possesses, is only now beginning to be understood. Second, we have little empirical data as to what classes of deductions are the most frequent in sophisticated language processing. It may be that special deductive techniques for the most common classes of inferences, together with parallelism, can overcome the efficiency difficulties.

One of the rules of inference the deductive mechanism will presumably provide is modus ponens. Thus, in our example, from

$$(\exists b) \, book(b)$$

and

$$(\forall x)(\exists y) \, book(x) \rightarrow index(y, x)$$

we will be able to conclude

$$(\exists b, y) \, book(b) \, \& \, index(y, b)$$

5. **Discourse Operations,** or **Specification of Possible Interpretations:** The deduction mechanism must be constrained, for we are able to infer much, much more from the propositional content of a sentence than is ever relevant. The approach I have taken is to say that a discourse presents us with certain "discourse problems", problems like coreference resolution, problems that must be solved if we are to be said to have understood the text. What counts as a solution can be specified in terms of inferences that can be drawn by the deductive mechanism from the propositional content of the sentence and the knowledge base. A possible interpretation of a sentence is taken to be a consistent combination of individual solutions to all of the sentence's discourse problems. The inferences that are relevant are then exactly those required by the "best" interpretation of the sentence, with "best" understood as explicated below.

We must therefore identify the discourse problems and, for each of them, specify what would count as a solution in terms of possible inferences. Somewhat more detail on this is presented below, but first let us consider our

5

example. One discourse problem is the problem of discovering the referent of a definite noun phrase, such as "the index" in sentence (3b). A solution might be specified in approximately the following manner:

> The existence of an entity of the description given by the definite noun phrase can be inferred from the previous text and the knowledge base, and that entity is the referent of the definite noun phrase.

Thus, because the deductive mechanism using modus ponens, as in (6), can infer from the expression (4a) in the representation of the previous text and from axiom (5) in the knowledge base that an index of book b exists, we assume that i is that index, thereby identifying z with b. The representation of text (3) now includes

$$(\exists i, b) \ldots \& \, book(b) \, \& \ldots \& \, index(i, b) \, \& \ldots$$

6. **Specification of the Best Interpretation:** The discourse operations only specify *possible* solutions to discourse problems, and there may be many. For example, in text (3) we may have the solution

> i is the index of book b,

or the solution

> i is the index of the first book listed in the bibliography of book b.

It is the task of Subtheory 6 to tell us that the first of these solutions is better than the second.

There has been very little work on this problem, although Nunberg (1978) has made a number of suggestions that deserve to be pursued. The basic idea is that we want to choose the most economical interpretation for the sentence as a whole. Among the factors that count in determining economy are the complexity of the proofs supporting the solutions, the salience of the axioms used, and certain redundancy properties in the interpretation.

Let us now look more closely at the discourse operations, closely enough to say just what the discourse problems are. In doing so, I would like to tell a story that suggests some logical necessity for just this set of discourse problems. We can divide the problems into those that arise in single sentences (whether or not they can be solved solely with information in the

sentence) and those that involve the relation of the sentence to something in the surrounding context.

Within the sentence: The logical form of a sentence consists of some logical combination of atomic predications, and an atomic predication consists of a predicate applied to one or more arguments. This suggests the following four classes of problems:

1. What does each argument refer to? This is the coreference resolution problem; it includes the subproblems of resolving pronouns, definite noun phrases, and missing arguments. In addition, many problems of syntactic ambiguity can be translated into coreference problems (Hobbs, 1982; Bear and Hobbs, 1988).

2. Where the predicate is nonspecific, what predicate is really intended? Examples of this problem are seen in compound nominals, denominal verbs, and uses of the possessive, the verb "have", and the prepositions "of" and "in".

3. How are the predicate and its arguments congruent? In the simplest cases this just involves the satisfaction of selectional constraints. When selectional constraints are not satisfied, there are two interpretive moves we can make. We can decide that the intended argument is not the explicit argument but something functionally related to it; this is *metonymy*. Or we can decide that the predicate does not mean what it ordinarily means, in the sense that some of the inferences one could ordinarily draw from its use are not appropriate in this instance; one example of this is *metaphor*.

4. Syntax tells us the *logical* relations among the atomic predications in the sentence, but frequently more information is conveyed. Consider for example the sentences

(4)   A car hit a jogger in Palo Alto last night.
(5)   A car hit a professor in Palo Alto last night.

Part of what is conveyed by sentence (4) is a causal relationship between the jogging and being hit by the car; inferring this relationship is essential to interpreting the sentence. We might call this the problem of determining *the internal coherence of the sentence.* Donnellan's (1966) referential-attributive distinction can be understood in these terms.

Beyond the sentence: Next we can ask what the relation is between the sentence and the surrounding environment (the "world"). In more operational terms, what is the relation between the logical form of the sentence and some internal representation of the environment? This is of course such

a huge problem it is certainly intractable. But there has been a great deal of work done in artificial intelligence on representing some aspects of the world as "plans" and attempting to specify how utterances relate to these plans. Such a plan may be a task model for some task the speaker and listener are executing jointly (Grosz, 1977; Linde and Goguen, 1978); it may be simply the speaker's presumed plan that led him to speak the utterance (Allen, 1979); it may be the listener's own conversational plan (Hobbs and Evans, 1980); or it may be the plan of a character in a story that is being told (Bruce and Newman, 1978; Wilensky, 1978). We might call all of this the problem of determining the *global coherence* of the utterance.

One of the most important things that is going on in the environment is the discourse itself. It is important enough to be singled out for special attention. The listener, in interpreting the sentence, must determine, consciously or subconsciously, its relation to the surrounding discourse. We might call this the problem of determining the *local coherence* of the utterance. It is on this problem that the remainder of the paper will be focused.

3   The Coherence Relations

The fundamental question that must be asked about discourse is, why is any discourse longer than one sentence? That is, why do we want to call a sequence of utterances a single discourse rather than simply a sequence of utterances? What are the definitional criteria for discourse?

We may approach the problem by describing as follows the situation in which discourse between a speaker and a listener takes place. (1) The speaker wants to convey a message. (2) The message is in service of some goal. (3) The speaker must link what he says to what the listener already knows. (4) The speaker should ease the listener's difficulties in comprehension.

These considerations give rise to four classes of coherence relations. In this section I take up each of the classes in turn. For each, the coherence relations in the class are motivated by the requirements of the discourse situation. A formal definition is given for each coherence relation in terms of the inferences a listener must draw, and a number of examples, together with the relevant inferences, are given. The examples are drawn from a wide variety of sources, including an algorithm description (Hobbs, 1977), a paragraph from *Newsweek* during the Watergate era (Hobbs, 1976), a life-history interview with a heroin addict (Agar and Hobbs, 1982), a medical

8

textbook on hepatitis, a book in archaeology, and several other sources.

There are two places in the discussion of the examples where I may seem to be appealing to magic. I often pull facts out of the hat, saying I am pulling them out of the knowledge base; and for every plausible analysis I present, I conceal a host of other analyses that cannot be ruled out by the definitions I give. Subtheory 3 of Section 2 allows me to pull the first of these tricks, while Subtheory 6 allows me to pull the second. Thus, whether the tricks are indeed magic remains to be seen, but they are, at the very least, beyond the scope of this chapter.

1. Frequently a message is coherent because it tells about coherent events in the world. It may seem that this observation converts a hard problem into an impossible one; instead of asking what makes a sequence of sentences in a text coherent, we ask what makes a sequence of *events* in the *world* coherent. But there are a few things we can say for certain about coherence in the world. First, temporal succession is not enough. We are often puzzled by two consecutive events if we can figure out no other relation between them than mere succession, and the same is true of two sentences in a discourse:

(6)   At 5:00 a train arrived in Chicago.
      At 6:00 Ronald Reagan held a press conference.

We may be able to read enough into the text to make it seem coherent, but it doesn't wear its coherence on its sleeve. When we start making assumptions to give it coherence, what criteria are we seeking to satisfy by means of the assumptions?

If we are able to see causality in the text, we are willing to conclude it is coherent. So if there is something special about the train—the maiden voyage of America's first bullet train, for example—to cause Reagan to call a press conference, then the text is coherent. But causality is too strong a requirement in general. Another way of reading (6) as coherent is by assuming that Ronald Reagan was on the train and the press conference was in Chicago. In this case there is no causal relation between the two events. It is a much weaker relation, one we might call an "occasion" relation, i.e., the first event sets up the occasion for the second.

The first coherence relation is thus the *occasion* relation. There are two cases, which may be defined as in (7). In this and in all the definitions we let S_1 be the current clause or larger segment of discourse, and S_0 an immediately preceding segment. For most of the examples we may assume the "assertion" of a clause to be what is predicated by the main verb; in

Type	(8a)	(8b)	(8c)
1	loc1 → loc2	loc2	
2		loc2	loc2 → loc3
2	angle1	angle1 → angle2	
1		angle1 → angle2	angle2

Figure 2: Occasion Relations in Example (8).

Section 4 there is some further discussion about what it is that segments assert.

(7) **Occasion:**
  1. A change of state can be inferred from the assertion of S_0, whose final state can be inferred from S_1.
  2. A change of state can be inferred from the assertion of S_1, whose initial state can be inferred from S_0.

Several instances of this relation occur in the following example from a set of directions:

(8a)   Walk out the door of this building.
(8b)   Turn left.
(8c)   Go to the corner.

Sentence (8a) describes a change of location whose final state holds during the event described in (8b). That location is the initial state in the change of location described in (8c). Similarly, an orientation is assumed in (8a) that is the initial state in a change of orientation described in (8b), and the final state of that change is assumed in (8c). There are thus four examples of the occasion relation in this text, as illustrated in Figure 2. Note that there is nothing wrong with finding more than one relation between sentences. If two relations do not involve inconsistent assumptions about indeterminate material in the text, there is no harm in saying that both relations obtain.

The following are further illustrations of the *occasion* relation and rough characterizations of the inferences that need to be drawn to satisfy the definition.

10

(9)   Decrease `N` by one.
      If it is zero, reset it to `MAX`.

The value of the variable `N` is changed, and the resulting value is presupposed in the second sentence.

(10)   He noticed the broken connection in the control mechanism,
       and took it to his workshop to fix.

The first clause asserts a change in knowledge that results in the action described in the second clause.

(11)   But they commonly doubt that the message is getting through to the President,
       and now their discouragement has been compounded by the news that Nixon's two savviest political hands, Melvin Laird and Bryce Harlow, plan to quit as soon as Ford settles in.

Discouragement being compounded is a change of mental state whose initial condition is the doubt described in the first clause.

(12)   But uh you know I dropped them [goods stolen from luggage] in my pocket,
       I tied the duffel bag up and the suitcase,
       and I left it there.

Dropping the goods frees the hands for tying, and the final state of the tying holds as the speaker leaves the luggage.

    Cause and *enablement* are important special cases of the *occasion* relation.

    2. The second class of coherence relations results from the need to relate what has been said to some goal of the conversation. I have called this *evaluation*. It can be defined as follows:

    Evaluation:
(13a)   From S_1 infer that S_0 is a step in a plan for achieving some goal of the discourse.

That is S_1 tells you why S_0 was said. The relation can also be reversed:

(13b)   From S_0 infer that S_1 is a step in a plan for achieving some goal of the discourse.

The discourse goal can be a very worldly goal, as in

> Did you bring your car today? My car is at the garage.

From the second sentence we can infer that the normal plan for getting somewhere in a car won't work, and that therefore the first sentence is a step in an alternate plan for achieving that goal.

Frequently, the goal is a conversational goal, for example, to entertain:

> The funniest thing happened to me.
> (A story).

or

> (A story).
> It was funny at the time.

It is because of this use that I have called this relation "evaluation". An important category of conversational goals is the goal of being understood.

> . . . Do you know what I mean?

Much "metatalk" is related to the rest of the discourse in this way.

This relation is close to the *cause* relation and to the *explanation* relation described below. If the state or event described in S_1 causes *the state or event described in S_0*, then S_1 *explains* S_0. If the state or event described in S_1 has caused *the speaker to say S_0*, then S_1 *evaluates* S_0.

3. The coherence relations in the third class are those directed toward relating a segment of discourse to the listener's prior knowledge. The two relations in this class are the *background* relation and the *explanation* relation.

First let us look at several examples of what one is inclined to call the background relation.

(14)   And one Sunday morning about ohhhh five o'clock in the morning I sat down in the Grand– no no, not in the Grand Central, in the Penn Station,
and while I was sitting there a young cat came up to me, . . .

(15)   In the round we were dancing I had barely noticed a tall, lovely, fair-haired girl they called Adrienne. All at once, in accordance with the rules of the dance, Adrienne and I found ourselves alone in the center of the circle. We were of the same height.
We were told to kiss and the dancing and the chorus whirled around us more quickly than ever.

(16)   `T` is a pointer to the root of a binary tree. ... The following algorithm visits all the nodes of the binary tree in inorder, making use of an auxilliary stack `A`.
T1: Initialize. Set stack `A` empty and set the link variable `P` to `T`.

It is not sufficient to say merely that the two segments refer to the same entities, for that would not rule out pairs like (17):

(17)   Ronald Reagan used to be a movie star.
He appointed George Shultz Secretary of State.

The first segment in each of the examples (14)-(16) seems to furnish background information for the second segment. It provides the "geography" against which the events of the second segment take place, or the "ground" against which the second segment places a "figure". But the "geography" can be quite metaphorical, as in example (16). Thus, a definition of the relation would be

Background:
Infer from S_0 a description of a system of entities and relations, and infer from S_1 that some entity is placed or moves against that system as a background.

This relation can occur in reverse order also, with the figure coming before the ground. This relation is of interest generally for causal reasons, for entities are causally influenced by the background against which they operate.
    The second relation in this class is *explanation*. Its definition is as follows:

(18)   **Explanation:**
Infer that the state or event asserted by S_1 causes or could cause the state or event asserted by S_0.

We don't need the inverse relation since we already have the relation *cause.* The *explanation* relation is a reason for telling a story backwards.

The following is a double example:

(19a)   He was in a foul humor.
(19b)   He hadn't slept well that night.
(19c)   His electric blanket hadn't worked.

Sentence (19b) tells the cause of the state described in (19a), while sentence (19c) gives us the cause of (19b). In the next example,

(20)   I thought well, maybe I can bum enough to get a cup of coffee and get into a movie,
'cause I was *exhausted*, I mean *exhausted*. My junk was running out.

the causality is explicitly indicated. But we would want to verify that the content is in accord with this. Exhaustion is a good reason to want shelter and, at least in the narrator's world, a movie theater is shelter. Finally, consider

(21a)   I said, hey look you guys, why don't you just soft-pedal it.
(21b)   I said, I don't know what your story is and I care less, but you're making a general display of yourself. This place is loaded with rats. It's only a matter of time until a cop comes in and busts the whole table.

The possible undesirable consequences described in (21b) are a cause for the behavior urged in (21a).

4. The final class of coherence relations, the "expansion" relations, is the largest. These are relations that, in a sense, expand the discourse in place, rather than carrying it forward or filling in background. They all involve inferential relations between segments of the text and can probably be thought of as easing the listener's inference processes. They can be classified in terms of moves between specific and general assertions and the interaction of these moves with negation, as illustrated in Figure 3. I have left two blank spaces in the "Negative" row because such relations would constitute a contradiction. They might be filled in with an "*exception*" relation. One states a general truth and then gives a specific exception to

	Specific to Specific	Specific to General	General to Specific
Positive:	Parallel	Generalization	Exemplification
Negative:	Contrast	–	–

Figure 3: The Expansion Coherence Relations.

it, or vice versa. But I have chosen rather arbitrarily to consider these as examples of *contrast*.

There are two important limiting cases. The *elaboration* relation is a limiting case of the *parallel* relation; the *violated expectation* relation is a limiting case of *contrast*.

Let us consider each of these relations in turn.

Parallel: The definition of the parallel relation is as follows:

(22) **Parallel:**
Infer $p(a_1, a_2, \ldots)$ from the assertion of S_0 and $p(b_1, b_2, \ldots)$ from the assertion of S_1, where a_i and b_i are similar, for all i.

Two entities are *similar* if they share some (reasonably specific) property. Determinations of similarity are subject to the same fuzziness and considerations of "good-ness" (Subtheory 6 again) as the coherence relations in general.

A simple example is this sentence from an algorithm description:

(23)   Set stack `A` empty and set link variable `P` to `T`.

From each of the clauses one can infer (trivially) that a data structure is being set to a value. The predicate p is thus *set*, stack `A` and link variable `P` are similar in that they are both data structures, and the stack's emptiness and `P`'s being equal to `T` are both initial conditions.

The next example is a bit more indirect. It comes from a problem in a physics textbook.

(24)   The ladder weighs 100 lb with its center of gravity 20 ft from the
       foot,
       and a 150 lb man is 10 ft from the top.

Because of the nature of the task, the reader must draw inferences from this
sentence about the relevant forces. We might represent the inferences as
follows:

(25) $force(100 \text{ lb}, L, Down, x_1) \,\&\, distance(F, x_1, 20 \text{ ft}) \,\&\, foot(F, L)$
       $force(150 \text{ lb}, x, Down, x_2) \,\&\, distance(T, x_2, 10 \text{ ft}) \,\&\, top(T, y)$

Here the predicate p is $force$, the first arguments are similar in that they
are both weights, the second and third arguments are both identical (once
we identify x with L), hence similar, and the fourth arguments are similar
in that they are points on the ladder at certain distances from an end of the
ladder (assuming y is L).
    The next example is from a medical textbook on hepatitis:

(26)   Blood probably contains the highest concentration of hepatitis
       B virus of any tissue except liver.
       Semen, vaginal secretions, and menstrual blood contain the agent
       and are infective.
       Saliva has lower concentrations than blood, and even hepatitis
       B surface antigen may be detectable in no more than half of
       infected individuals.
       Urine contains low concentrations at any given time.

The predicate p is $contain$; the diagram in Figure 4 indicates the correspond-
ing similar arguments and the shared properties (the column headings) by
virtue of which they are similar.
    The next example is from Shakespeare's 64th sonnet:

(27)   When sometime lofty towers I see down-rased
       And brass eternal slave to mortal rage;

We would like to understand the chain of inferences that establish the *parallel*
relation between "sometime lofty towers ... down-rased" and "brass eternal
slave to mortal rage." From "down-rased" we can infer that the towers are

BODY MATERIAL	CONTAINS	CONCENTRATION	AGENT
blood	contains	highest concentration	HBV
semen vaginal secretions menstrual blood	contain		agent
saliva	has	lower concentrations	
(saliva of) infected individuals	in	detectable … no more than half	HBsAg
urine	contains	low concentrations	

Figure 4: The *Parallel* Relation in Example (26).

destroyed. There are several possible interpretations of "mortal rage", but one is that mortal rage is death. To be slave to death is to be controlled by death, and thus to be destroyed. Therefore, the predicate p which each half of the parallelism asserts is *destroyed*. Next it must be determined in what way lofty towers and brass eternal are similar. Towers, being buildings, are (relatively) permanent. Brass, being metal, is relatively permanent, and of course "eternal" implies "permanent" directly. Thus towers and brass are similar in that they are at least seemingly permanent. These clauses are interesting also because they have an internal coherence relation of violated expectation: seemingly permanent entities are destroyed.

The next example is a Congressman's complaint about communication with the Nixon White House staff, quoted in the *Newsweek* paragraph:

(28)   We have nothing to say to Ron Ziegler,
       and Al Haig's never been in politics.

The *parallel* relation here depends on the inference from each clause that Ron Ziegler and Al Haig (similar entities, in that both were advisors to Nixon) are people with whom members of Congress cannot communicate.

Finally, an example from the heroin addict's life history:

(29) But he *had* a really *fine* pair of *gloves*,
      and uh along with the gloves he had uh a– a cheap camera, I don't
      know, it was a– a *Brownie*, I think,
      and one or two other little objects that didn't amount to doodly
      doo.

The three clauses are in a *parallel* relation because each asserts the existence
and expresses an evaluation of objects in stolen luggage.

Elaboration: The *elaboration* coherence relation is just the *parallel*
relation when the similar entities a_i and b_i are in fact identical, for all i. It
can be given the following definition:

(30) **Elaboration:**
      Infer the same proposition P from the assertions of S_0 and S_1.

Frequently the second segment adds crucial information, but this is not
specified in the definition since it is desirable to include pure repetitions
under the heading of *elaboration*.

A simple illustration of the *elaboration* relation is the following:

(31) Go down First Street.
      Just follow First Street three blocks to A Street.

From the first sentence we can infer

(32a) *go*(`Agent`: you, `Goal`: x, `Path`: First St., `Measure`: y)

for some x and y. From the second we can infer

(32b) *go*(`Agent`: you, `Goal`: A St., `Path`: First St., `Measure`: 3 blks)

If we assume that x is A Street and y is 3 blocks, then the two are identical
and serve as the proposition P in the definition.

A slightly more interesting case is

(33) John can open Bill's safe.
      He knows the combination.

From the first sentence and from what we know about "can", we can infer that John knows some action that will cause the safe to be open. From the second sentence and from what we know about combinations and knowledge, we can infer that *he*, whoever *he* is, knows that dialing the combination on whatever it is the combination of will cause it to be open. By assuming that "he" refers to John and that the combination is the combination of Bill's safe, we have the same proposition P and have thus established the *elaboration* relation (and solved some coreference problems as a by-product—see Hobbs, 1980).

The next example is from a book on the archaeology of China:

(34)　This immense tract of time is only sparsely illuminated by human relics.
Not enough material has yet been found for us to trace the technical evolution of East Asia.

From "sparse" and "illuminate" we can infer in the first sentence that the relics fail to cause one to know the "contents" of the immense tract of time. From "not enough" in the second sentence, we can infer that the material fails to cause us to know the "contents" of the technical evolution. "Relics" and "material" are the same, as are the "immense tract of time" and "the technical evolution of East Asia". The proposition P is therefore something like "The material found does not cause us to know the contents of a tract of time."

The next example is from the medical text:

(35)　Generally blood donor quality is held high by avoiding commercial donors . . .
Extremely careful selection of paid donors may provide safe blood sources in some extraordinary instances, but generally it is much safer to avoid commercially obtained blood.

Here it is crucial to recognize that blood donor quality being held high is a way of minimizing risk, which implies greater safety.

Another from the *Newsweek* paragraph:

(36)　Time is running out on Operation Candor.
Nixon must clear himself by early in the new year or lose his slipping hold on the party.

Recognition of the *elaboration* relation depends on inferring the commonality between "time is running out on ..." and "must ... by early in the new year", and then recognizing, either by knowing or assuming, that "Operation Candor" and "Nixon ... clear himself" are identical complexes of events.

Finally:

(36)   Al Haig's never been in politics–
        he can't even spell the word "vote."

Both clauses are intended to imply that Haig is not knowledgeable about politics—the first by saying that he lacks the relevant experience, the second by giving an alleged example of some "political" skill he lacks.

For simplicity, in the remaining definitions, it will be assumed that the assertions of the segments that the relation links are predications with one argument. The definitions can be extended in a straightforward manner to more than one argument.

Exemplification: The *exemplification* relation is defined as follows:

(38)   **Exemplification:**
        Infer $p(A)$ from the assertion of S_0 and $p(a)$ from the assertion of S_1, where a is a member or subset of A.

A fairly simple example is the following:

(39)   This algorithm reverses a list.
        If its input is "A B C", its output is "C B A".

Recognizing the relation depends on inferring "causes X to be the reverse of X" from "reverses", inferring the causal relation between the input and output of an algorithm, recognizing that "A B C" is a list and that "C B A" is its reverse.

This more complex example is from the archaeology text:

(40)   We cannot affirm that the technical evolution of East Asia followed
        the same course as it did in the West.
        Certainly no stage corresponding to the Mousterian tradition has
        been found in China.

"Cannot affirm" is matched by "no stage ... has been found". China is a part of East Asia, and "stage ... in China" is one portion of "the technical evolution of East Asia", just as the Mousterian tradition is a portion of the technical evolution of the West.

The *generalization* coherence relation is simply *exemplification* with S_0 and S_1 reversed.

Contrast: There are two cases of the *contrast* relation. They can be characterized as follows:

Contrast:

(41a)   Infer $p(a)$ from the assertion of S_0 and $\neg p(b)$ from the assertion of S_1, where a and b are similar.

(41b)   Infer $p(a)$ from the assertion of S_0 and $p(b)$ from the assertion of S_1, where there is some property q such that $q(a)$ and $\neg q(b)$.

In the first case, contrasting predications are made about similar entities. In the second case, the same predication is made about contrasting entities.

The first example illustrates the first case:

(42)   You are not likely to hit the bull's eye,
       but you are more likely to hit the bull's eye than any other equal
       area.

From the first clause we can infer that the probability of hitting the
bull's eye is less than whatever probability counts as likely. From the second clause we can infer that the probability is greater than (and thus not less than) the typical probability of hitting any other equal area.

The second example illustrates the second case:

(43)   If `INFO(M) > INFO(N)`, then set `M` to `LINK(M)`.
       If `INFO(M) ≤ INFO(N)`, then set `N` to `LINK(N)`.

What is asserted in each sentence is an implication. The first arguments of the implications are contradictory conditions. The second arguments are similar in that they are both assignment statements. Note that we must discover this relation in order not to view the instructions as temporally ordered and thereby translate them into the wrong code.

Finally, consider

(44)   Research proper brings into play clockwork-like mechanisms;
        discovery has a magical essence.

"Research" and "discovery" are viewed as similar elements, "mechanistic" and "magical" as being contradictory. This therefore illustrates the first case.

Violated Expectation: The final coherence relation is the *violated expectation* relation, defined as follows:

(45)   **Violated Expectation:**
        Infer P from the assertion of S_0 and $\neg P$ from the assertion of S_1.

    An example would be

(46)   John is a lawyer, but he's honest.

Here one would draw the inference from the first clause that John is dishonest since he is a lawyer, but that is directly contradicted and thus overridden by the second clause.

    In the following sentence from a referee's review,

(47)   This paper is weak, but interesting.

one can infer from the first clause that the paper should be rejected, but from the second clause that it should be accepted.

    Next:

(48)   The conviction is widespread among Republicans that Mr. Nixon
        must clear himself by early in the new year.
        But they commonly doubt that the message is getting through to
        the President.

Typically, if something is true of a person, that person would be expected to know it. But the second sentence denies that.

    The final and most complex example is from Lenin's *State and Revolution*.

22

(49)   We are in favor of a democratic republic as the best form of the
       state for the proletariat under capitalism;
       but we have no right to forget that wage slavery is the lot of the
       people even in the most democratic bourgeois republic.

The democratic republic is best for the people under capitalism, but contrary
to what one might expect from this, a rather undesirable condition—wage
slavery—would still obtain.

From one perspective, we can view the coherence relations as text-
building strategies, strategies the speaker uses to make the listener's com-
prehension easier. But that does not answer the question of why this partic-
ular set of relations should make comprehension any easier. It is tempting
to speculate that these coherence relations are instantiations in discourse
comprehension of more general principles of coherence that we apply in at-
tempting to make sense out of the world we find ourselves in, principles
that rest ultimately on some notion of cognitive economy. We get a simpler
theory of the world if we can minimize the number of entities by identifying
apparently distinct entities as different aspects of the same thing. Just as
when we see two parts of a branch of a tree occluded in the middle and
assume that they are parts of the same branch, so in the expansion relations
we assume that two segments of text are making roughly the same kind of
assertion about the same entities or classes of entities. When we hear a
loud crash and the lights go out, we are apt to assume that one event has
happened rather than two, by hypothesizing a causal relation. Similarly, the
weak sort of causality underlying the *occasion* relation seems to be a way of
binding two states or events into one. Recognizing coherence relations may
thus be just one way of using certain very general principles for simplifying
our view of the world.

4   The Structure of Discourse

A clause is a segment of discourse, and when two segments of discourse
are discovered to be linked by some coherence relation, the two together
thereby constitute a single segment of discourse. By recognizing coherence
relations between segments, we can thus build up recursively a structure for
the discourse as a whole. For example, clauses (1b) and (1c) in text (1) are

linked by an *occasion* relation. They combine into a segment that is in turn related to clause (1a) by an *elaboration* relation. This results in a composed segment that consists of all of sentence (1); this is related to sentence (2) by an *occasion* relation. We can call the resulting structure for the text its "coherence strucure". Typically, in a well-organized written text, there will be one tree spanning the entire discourse.

This notion of structure in discourse allows us to get a handle on some classical problems of discourse analysis. Here I will touch on just three: the notion of "topic", one aspect of the notion of "genre", and some of the deviations from coherence that occur in ordinary conversation.

There are really two notions of "topic" (and I refer here and throughout only to discourse topic, not sentence topic). A topic is a segment of a discourse about a single thing, and a topic is a characterization of the thing a segment is about. The first notion of topic is easy to characterize in terms of the coherence structure of texts. It is a segment spanned by a single tree which is not included in a larger segment spanned by a single tree.

There may seem to be problems with this definition when topic boundaries are uncertain. In a dialog analyzed by David Evans and me (Hobbs and Evans, 1980), there is a stretch of talk about the contents of envelopes the woman is carrying, and then about her dissertation, a copy of which she is also carrying. Are there two topics—envelopes and dissertation—or just one—things she is carrying? It is hard to know what the maximal segments should be. But this uncertainty as to topic structure is exactly reflected in the uncertainty as to whether there is a *parallel* coherence relation between the two segments. Is the fact that she is carrying both the envelopes and the dissertation sufficient for the similarity required by the definition of the *parallel* relation?

The problem of characterizing the second notion of topic is a bit more difficult, and we need to back up and discuss another problem that has heretofore been glossed over. The definitions of the coherence relations are stated in terms of what utterances *assert*. In many cases it is simple to decide what is asserted: the predication expressed by the main verb. So in

The boy hit the ball.

we are asserting something like $hit(BOY_1, BALL_1)$. But there are many utterances in which this simple rule does not apply. In

They hanged an innocent man today.

24

it may already be mutually known that they hanged someone, and the speaker is asserting the man's innocence. This is a serious problem, but I do not have space to address it here. For the purposes of discussion let us assume that the assertions of single utterances can be determined.

If the definitions of the coherence relations can be applied to segments of discourse larger than a single clause, we need to be able to say what is asserted by those segments. We can do so if, in the composition process, when two segments S_0 and S_1 are joined by a coherence relation into a larger segment S, we have a way of assigning an assertion to S in terms of the assertions of S_0 and S_1. The assertion of S will consititute a kind of summary of the segment S.

As an approach to this problem we can divide the relations into two categories: coordinating and subordinating. Among the coordinating relations are *parallel* and *elaboration*. To recognize a coordinating relation, one must generally discover some common proposition inferable from each segment. We can assign this common proposition as the assertion of the composed segment. For the *parallel* relation, we must infer $p(a)$ and $p(b)$, where a and b are similar by virtue of sharing some property q. We can then say that the composed segment asserts $p(x)$ where x is in $\{x|q(x)\}$. For example, in (24) the assertion of the whole is something like "There are downward forces acting on the ladder at some distance from an end of the ladder." In (27) the assertion is "Seemingly permanent things are destroyed." For the *elaboration* relation, we must infer some proposition P from the assertion of each segment. We can say that P is the assertion of the composed segment. In (33) the assertion is that John knows that dialing the combination will cause the safe to be open.

Among the subordinating relations are *background*, *explanation*, *exemplification* and *generalization*, *contrast*, and *violated expectation*. In these relations one of the two segments, S_0 or S_1, is subordinated to the other. We can say that the assertion of the composed segment is the assertion of the dominant segment. In the *contrast* and *violated expectation* relations "S_0, but S_1", it is generally the second segment that is dominant, although there are exceptions. Thus, sentence (47) urges acceptance of the paper (I am happy to report). In *exemplification* and *generalization*, it is the more general statement. In *explanation* "S_0, because S_1", it is the first segment. For the *background* relation, the dominant segment is the figure, the segment for which the background is provided.

I'm not sure what to say about the *occasion* relation, whether to say that the composed segment asserts the assertion of the second segment, that it

asserts the change, or that it asserts the occurrence of some abstract event which decomposes into the stated events.

With rules such as these for assigning assertions to larger segments of discourse, it becomes easy to define the second notion of topic. A topic-in-the-first-sense is a composed segment. The topic-in-the-second-sense of this segment is the assertion assigned to it by the above rules, i.e., a kind of summary of its contents.

With this notion of discourse structure we can begin to examine conventional structures peculiar to certain genres. There are in principle many ways one could structure an account of a sequence of events, but in a given genre, for one reason or another, a few of the ways have been institutionalized or conventionalized into frozen forms. It is these constrained coherence structures that researchers who propose story grammars are seeking to characterize.

As an illustration, let us look at a conventional coherence structure for narratives that, to my knowledge, has not previously been observed. It is exhibited in the following two stories. The first is from the life-history interviews with the heroin addict:

(50a)   And one Sunday morning about ohhhh five o'clock in the morning I sat down in the Grand– no no, not in the Grand Central, in the Penn Station,

(50b)   and while I was sitting there a young cat came up to me, and he had his *duffel* bag and a suitcase, and he said, "Look," he said, "maaan," he said, "I've got to make the john. Will you keep your *eye* on the– on my stuff for me?"
Well there were two . . black fellows sitting down at the end of the line, watching this procedure, you know and I –

(50c)   for a few minutes I thought well fuck it, I – you know I'm gonna – the guy trusts me, what's the use of trying to beat him.

(50d)   But one of the black guys came over, and said, "Hey maaan, why don't you dig in and see what's there, maaan, maaan, you know, maybe we can split it,"

(50e)   and I said we're not going to split it at all, it's mine, and I picked
         up the suitcase, threw the duffel bag over my back and I *split*,

(50f)   and left a very irritated guy there, "I'll catch you motherfucker," he
         said, and I said, "well maybe you will and maybe you won't," and
         I'm hightailing it as fast as I can.

The second is from a life story collected by Charlotte Linde (Linde, 1984).

(51a)   Uh, I started out in Renaissance studies,

(51b)   but I didn't like any of the people I was working with,

(51c)   and at first I thought I would just leave Y and go to another
         university,

(51d)   but a medievalist at Y University asked me to stay or at least re-
         consider whether I should leave or not, and um pointed out to me
         that I had done very well in the medieval course that I took with
         him and that I seemed to like it, and he was right. I did.

(51e)   And he suggested that I switch fields and stay at Y

(51f)   and that's how I got into medieval literature.

Both have the structure illustrated in Figure 5.

In each story, segment (a) provides background for (b). The circum-
stance of segment (d) causes and thus occasions the events of (e). Segments
(c) and (d)-(e) are contrasting solutions. Segments (a)-(b) and (c)-(e) are
related by an important subtype of the *occasion* relation—a problem and
its solution. Segments (a)-(e) and (f) are related by another important
subtype—a set of events and its outcome.

It is likely that this structure is a very common pattern for stories in our
culture. It is a coherence structure, but not just any coherence structure.
In this convention, the *occasion* relations are constrained to be a problem-
solution relation and an event-outcome relation, and the *contrast* has to be
between two possible solutions.

Other genres have similar conventional constrained coherence structures.
Considerations of coherence in general allow us to string together arbitrar-
ily many parallel arguments. But it is a convention of argumentation for
there to be just three, and those ordered by increasing strength. In political
rhetoric, one also hears sequences of parallel statements, but for maximum
effectiveness, they should be more than just the semantic parallelisms char-
acterized by the theory of coherence. They should also exhibit a high degree
of lexical and syntactic parallelism.

```
                              Occasion
                             /        \
                    Occasion          Outcome
                   /        \             (f)
          Problem:          Solutions:
   Background → Statement     Contrast
        /         \          /        \
   Setting      Problem    First      Second
     (a)          (b)   Alternative  Alternative:
                           (c)      Cause → Effect
                                      /        \
                               Circumstance    Second
                                    (d)      Alternative
                                                 (e)
```
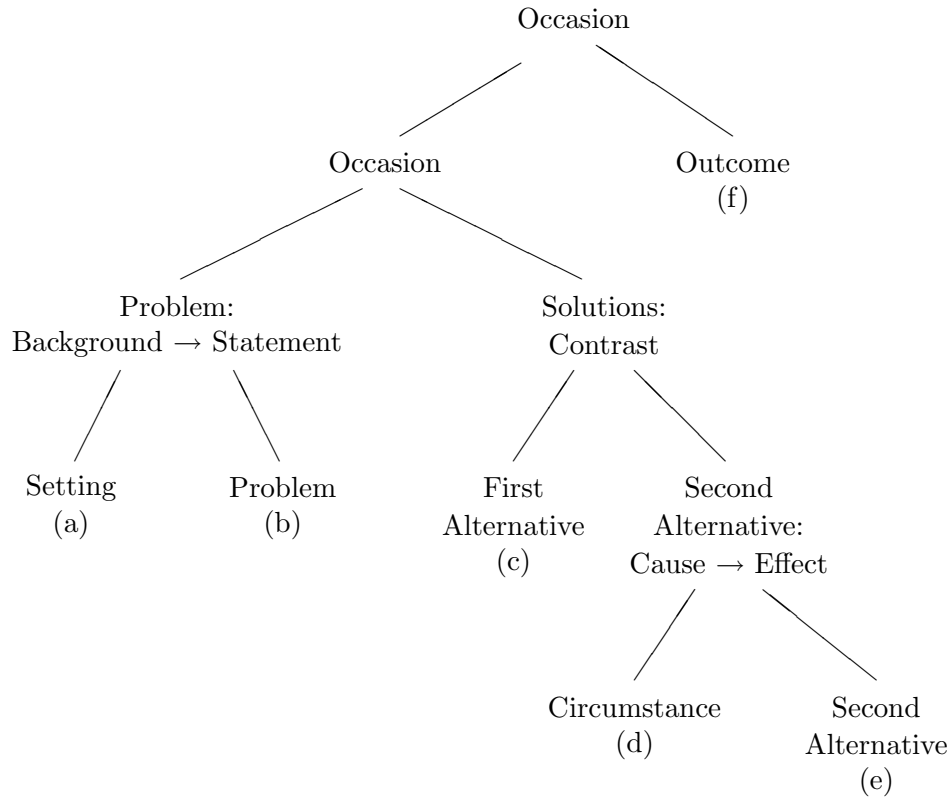
Figure 5: The Structure of Stories (50) and (51).

In a well-planned text, it is possible that one tree will span the entire text. However, conversations drift. We are likely to see a sequence of trees spanning conversational segments of various sizes, with perhaps smaller trees spanning the gaps between the larger segments—something resembling what is shown in Figure 6. To switch metaphors in midforest, we see a number of more or less large islands of coherence linked by bridges of coherence between two points at the edges of the islands. Thus, the first sentence of a new island may be in a parallel relation to the last sentence of the previous island, but in a way that fails to develop the structure of either island.

A notorious example of such local coherence and global incoherence is the phenomenon of going off on a tangent. An example of this occurs in the life-history interviews between the heroin addict Jack and the anthropologist
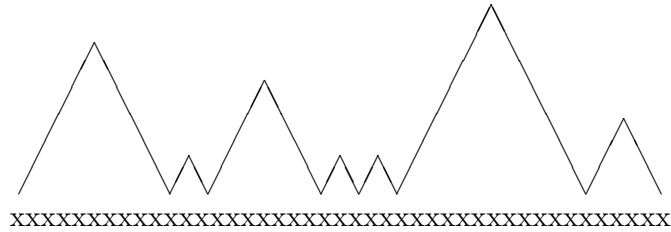
Figure 6: Typical Structure of a Conversation.

Mike Agar (Hobbs and Agar, 1984). This interview began with Agar asking and Jack agreeing to talk about Jack's move from Chicago to New York when he was fifteen. After explaining *why* he left Chicago, Jack is now telling *how* he did it—by hitchhiking. He mentions his previous experience with hitchhiking and then slides into a reminiscence about a trip to Idaho.

(52a)	J:	I had already as I told you learned a little bit about hitchhiking,
(52b)	J:	I'd split out and uh two or three times, then come back,
	M:	Uh huh.
(52c)	J:	The one – my first trip had been to Geneva uh New York,
	M:	Uh huh.
(52d)	J:	And then I'd uh once or twice gone to – twice I'd gone to California,
(52e)		And then I'd cut down through the South,
(52f)		And I had sort of covered the United States.
(52g)		One very beautiful summer I'll tell you about some other time that I spent in Idaho
(52h)		That to this day I remember with nothing but you know happiness,
(52i)		It was so beautiful,
(52j)		I'll- I'll never forget it,
(52k)		I- Right up in the mountains in these tall pine forests,
(52l)		And it was something that you know is just- it you know-
(52m)	J:	It's indelibly in my memory,
	M:	That's huh
(52n)	J:	And nothing could ever erase it.
(52o)	M:	We'll have to- we'll come back to it one day.

29

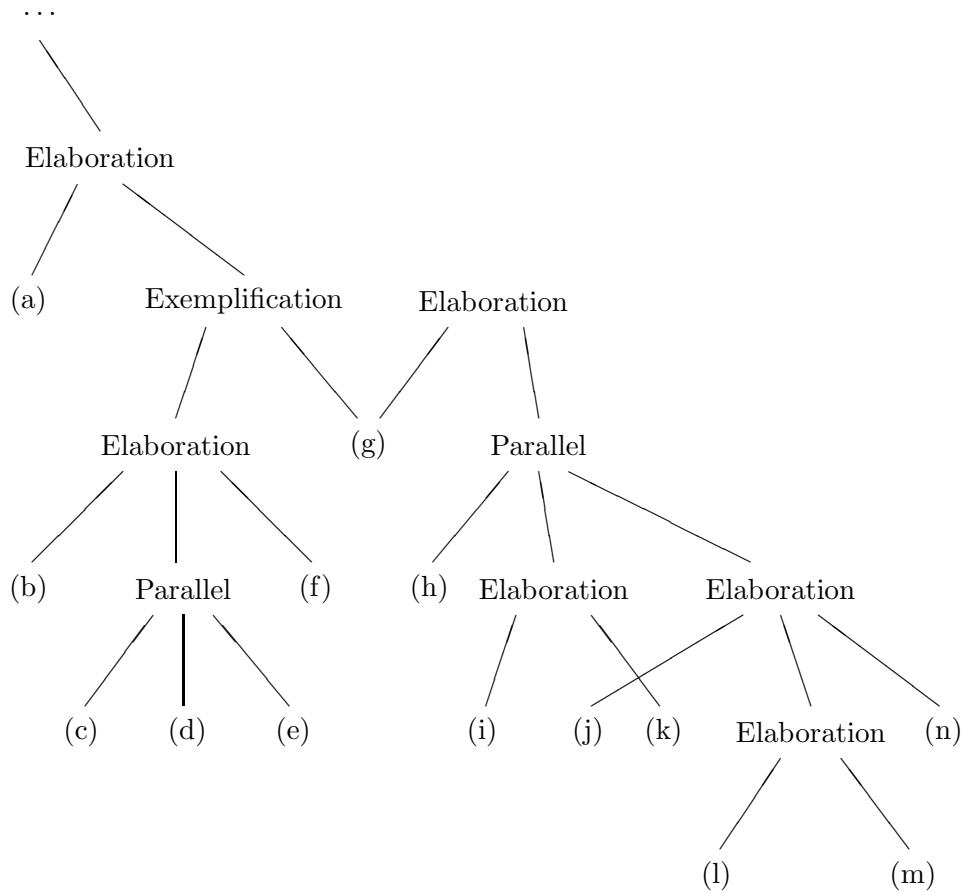Figure 7: Structure of Example (52).

(52p) J: Yeah, sometime you ask me about that.
(52q) M: Okay.

Figure 7 illustrates the structure of this passage.

In utterance (52a), Jack is working out a reasonable step in his global plan, namely, to explain that he had the means to leave Chicago—hitchhiking. He elaborates on this in (52b) to (52e) by giving several parallel examples of his experiences with hitchhiking, summing up in (52f). In (52g) he gives one final example, and here the tangent begins as he elaborates on the beauty of the summer. In (52h) he tells of his happiness. In (52i) he repeats that

it was beautiful. In (52j) he says he'll never forget it. In (52k) he gets specific about what was beautiful. Utterance (52l) is probably a false start for (52m), and in (52m) and (52n) he says again in two different ways that he'll never forget it.

It is interesting to see how this slide happens. The crucial utterance is (52g). Significantly, it is not clear whether it is a topicalized sentence or just a noun phrase. It is uncertain which predication is to be treated as its assertion. Insofar as it is an *exemplification* of (52f), the assertion is "I spent one summer in Idaho." But the predication that is elaborated upon subsequently, and thus functions as the assertion of (52g) from the perspective of the last half of the passage, is "The summer was very beautiful." It is the ambiguity in what (52g) asserts that enables the tangent to occur. The anthropologist finally redirects the interview in (52o) by picking up on the third predication made in (52g)—"I'll tell you about the summer some other time"—and the interview gets back on track.

This example suggests an enrichment of our view of the function of the coherence relations. The coherence relations are not merely constraints on the orderly top-down development of discourse. They are also resources to which the speaker may appeal to get him from one sentence to the next when global constraints are insufficient or insufficiently attended to. They are a means of finding a next thing to say (see also Hobbs and Evans, 1980). A tangent occurs when there is a kind of relaxation in the discourse planning process and local coherence is pursued to the neglect of global concerns.

5   A Method for Analyzing Discourse

This account of the structure of discourse suggests a method for analyzing discourse. The method consists of four steps, each an order of magnitude more difficult than the one before it.

1. One identifies the one or two major breaks in the text and cuts it there. That is, one chooses the most natural way to divide the text into two or three segments. This can be done on a strictly intuitive basis by anyone who has understood the text, and among those who have understood it in the same way, there will be a large measure of agreement. This process is then repeated for each of the segments, dividing *them* in the most natural places. The process is continued until reaching the level of single clauses. This yields a tree structure for the text as a whole.

In text (1), for example, the major break comes between sentences (1)

and (2). Within sentence (1) there is a break between the first clause and the last two, and of course a final break betwen the second and third clauses of the first sentence. This yields the tree of Figure 1.

2. One labels the nonterminal nodes of the tree with coherence relations. Proceeding from the bottom up, one devises rough accounts of what is asserted by each composed segment. Thus, in the Chomsky example, we label the node linking (1b) and (1c) with the *occasion* relation. We label the node linking the resulting segment and (1a) with the *elaboration* relation. Finally, we label the node linking (1) and (2) with the *occasion* relation.

In this step the method becomes theory-specific, as one must know what the relations are and have at least rough characterizations of them. One aid in this task is to determine what conjunctions or sentential adverbs it would be appropriate to insert. If we can insert "then" between S_0 and S_1, and the sense would be changed if we reversed the segments, then the *occasion* relation is an excellent candidate. If we can insert "because", the *explanation* relation becomes a strong possibility. "That is" or "i.e." suggests *elaboration*, "similarly" suggests *parallel*, "for example" suggests *exemplification*, and "but" suggests *contrast* or *violated expectation*. It should be emphasized, however, that these tests are informal. They do not define the relations. Conjunctions and sentential adverbials impose constraints on the propositional content of the clauses they link or modify, and in many cases these constraints are almost the same as those imposed by some coherence relation. In the best of cases there is sufficient overlap for the conjunction to tell us what the coherence relation is.

3. One makes (more or less) precise the knowledge or beliefs that support this assignment of coherence relations to the nodes. Each of the coherence relations has been defined in terms of the inferences that must be drawn from the listener's knowledge base in order to recognize the relation. When we say, for example, that an occasion relation occurs between (1b) and (1c), we have to specify the change asserted in (1b) (namely, a change in mutual knowledge about where the controversy lies, from the word "identify") that is presupposed in the event described in (1c), (the effort to resolve the controversy). Thus, we need knowledge about what change is effected by the action of identifying, and we need to know the meanings of "controversy" and "resolution" that allow us to talk about controversies being resolved.

The precision with which we specify the knowledge really can be "more or less". We might be satisfied with a careful statement in English, or we might demand formulation in terms of some logical language, embedded within a larger formal theory of the commonsense world.

4. One validates the hypotheses made in step 3 about what knowledge underlies the discourse. Mike Agar and I (Agar and Hobbs, 1982) have discussed at length how this should proceed. Briefly, one looks at the larger corpus to which the text belongs, a corpus by the same speaker or from the same culture that assumes the same audience. One attempts to construct a knowledge base or system of mutual beliefs that would support the analyses of all of the texts in the corpus. If step 1 is a matter of minutes for a text of paragraph length, step 2 a matter of an hour or two, and step 3 a matter of days, then step 4 is a matter of months or years.

In each of these steps difficulties may arise, but these difficulties in analysis will usually reveal problematic aspects of the text. In step 1, we might find it difficult to segment the text in certain places, but this probably reflects a genuine area of incoherence in the text itself. We might find it easy to segment the text because the segments are about clearly different topics, but be unable to think of a coherence relation that links the segments. When this happens, it may be that we have found two consecutive texts rather than a single text. At times the knowledge that underlies a composed segment is not obvious, but this often leads us to very interesting nonstandard assumptions about the belief systems of the participants. For example, to justify the *explanation* relation in (20), we have to assume it is mutually understood that movie theatres are shelters. Finally, we often cannot be sure the knowledge we have assumed to be operative really is operative; looking at further data forces revisions in our assumptions.

The theory of local coherence in discourse that I have sketched in this paper is part of a larger theory that seeks to make explicit the connection between the interpretation of a text and the knowledge or belief system that underlies the text. The coherence relations that give structure to a text are part of what an interpretation is; they are defined in terms of inferences that must be drawn to recognize them, and thus specify one connection that must exist between interpretations and knowledge. The method outlined in this section can be used to exploit that connection in several ways.

Where, as in ethnography, our interest is in the belief systems, or the culture, shared by the participants, the method acts as a "forcing function". It does not tell us what the underlying beliefs are, but it forces us to hypothesize beliefs we might otherwise overlook, and it places tight constraints on what the beliefs can be.

Where our interest is primarily in the interpretation of the text, as in literary criticism, the method gives us a technique for finding the structure of the text, an important aspect of the interpretation. In placing constraints

on the ideal structure of a text, it can point us toward problematic areas of the text where the ideal of coherence proposed here does not seem to be satisfied. We might ultimately decide in such cases that the ideal is in fact not satisfied, but many times we will find that the attempt to satisfy the ideal leads us to interesting reinterpretions of the whole text. The next chapter provides an example of this.

References

[1] Agar, Michael, and Jerry R. Hobbs, 1982. "Interpreting Discourse: Coherence and the Analysis of Ethnographic Interviews", *Discourse Processes*, Vol. 5, No. 1, pp. 1-32.

[2] Allen, James, 1979. "A Plan-based Approach to Speech Act Recognition", Technical Report No. 131/79, Department of Computer Science, University of Toronto, Canada.

[3] Bruce, Bertram C., and Dennis Newman, 1978. "Interacting Plans", *Cognitive Science*, Vol. 2, pp. 195-233.

[4] Crothers, Edward, 1979. *Paragraph Structure Inference*, Ablex Publishing Corporation, Norwood, New Jersey.

[5] Dahlgren, Kathleen, 1985. "The Cognition Structure of Social Categories", *Cognitive Science*, vol. 9, no. 3, July-September 1985, pp. 379-398.

[6] Donnellan, Keith S., 1966. "Reference and Definite Descriptions", *The Philosophical Review*, Vol. 75, pp. 281-304.

[7] Fillmore, Charles, 1974. "Pragmatics and the Description of Discourse", in *Berkeley Studies in Syntax and Semantics*, Vol. 1, C. Fillmore et al, eds., University of California, Berkeley, California, pp. V-1-V-21.

[8] Grimes, Joseph, 1975. *The Thread of Discourse*, Mouton and Company, The Hague, Netherlands.

[9] Grosz, Barbara, 1977. "The Representation and Use of Focus in Dialogue Understanding", Stanford Research Institute Technical Note 151, Stanford Research Institute, Menlo Park, California. July 1977.

[10] Hobbs, Jerry R., 1976. "A Computational Approach to Discourse Analysis", Research Report 76-2, Department of Computer Sciences, City College, City University of New York. December 1976.

[11] Hobbs, Jerry R., 1977. "From 'Well-Written' Algorithm Descriptions into Code", Research Report 77-1, Department of Computer

Sciences, City College, City University of New York. July 1977.

[12] Hobbs, Jerry R., 1979. "Coherence and Coreference", *Cognitive Science*, Vol. 3, No. 1, pp. 67-90.

[13] Hobbs, Jerry R., 1982. "Representing Ambiguity", *Proceedings of the First West Coast Conference on Formal Linguistics*, Stanford, California, January 1982, pp. 15-28.

[14] Hobbs, Jerry R., and Michael Agar, 1985. "The Coherence of Incoherent Discourse", *Journal of Language and Social Psychology*, vol. 4, nos. 3-4, pp. 213-232.

[15] Hobbs, Jerry R., William Croft, Todd Davies, Douglas Edwards, and Kenneth Laws, 1987. "Commonsense Metaphysics and Lexical Semantics", *Computational Linguistics*, vol. 13, nos. 3-4, July-December 1987, pp. 241-250.

[16] Hobbs, Jerry R., and David Andreoff Evans, 1980. "Conversation as Planned Behavior", *Cognitive Science*, Vol. 4, No. 4, pp. 349-377.

[17] Hobbs, Jerry R., and Robert C. Moore, eds., 1985. *Formal Theories*

of the Commonsense World, Ablex Publishing Corporation, Norwood, New Jersey.

[18] Hovy, Eduard H., 1988. "Planning Coherent Multisentential Text", *Proceedings of the 26th Annual Meeting, Association for Computational Linguistics*, pp. 163-169. Buffalo, New York, June 1988.

[19] Lenat, Doug, Mayank Prakash, and Mary Shepherd, 1986. "CYC: Using Common Sense Knowledge to Overcome Brittleness and Knowledge Acquisition Bottlenecks", *AI Magazine*, vol. 6, no. 4, Winter 1986, pp. 65-85.

[20] Linde, Charlotte, 1989. *The Creation of Coherence in Life Stories*, to be published by Ablex Publishing Corporation, Norwood, New Jersey.

[21] Linde, Charlotte, and Joseph Goguen, 1978. "Structure of Planning Discourse", *Journal of Social and Biological Structures*, Vol. 1, pp. 219-251.

[22] Longacre, Robert, 1976. *An Anatomy of Speech Notions*, The Peter de Ridder Press, Ghent, Belguim.

[23] Mann, William, and Sandra Thompson, 1986, "Relational Propositions in Discourse", *Discourse Processes*, Vol. 9, No. 1, pp. 57-90.

[24] Meyer, Bonnie, 1975. "Identification of the Structure of Prose and its Implications for the Study of Reading and Memory", *Journal of Reading Behavior*, Vol. VII, No. 1, pp. 7-47.

[25] Montague, Richard, 1974. "The Proper Treatment of Quantification in Ordinary English", in *Formal Philosophy: Selected Papers*, R. Thomason, ed., Yale University Press, New Haven, Connecticut.

[26] Nunberg, Geoff, 1978. "The Pragmatics of Reference", Ph.D. thesis. City University of New York, New York, New York.

[27] Wilensky, Robert, 1978. "Understanding Goal-Based Stories." Yale University Dept. of Computer Science Research Report 140. September 1978.

[28] Woods, William, 1970. "Transition Network Grammars for Natural Language Analysis", *Communications of the ACM*, Vol. 13, pp. 591-606.