DISCOURSE AND COHESION IN EXPOSITORY TEXT†

Allen B. Tucker Sergei Nirenburg* and Victor Raskin***

*Department of Computer Science, Colgate University
**Department of English, Purdue University

1. Background and Introduction

This paper discusses the role of discourse in expository text; text which typically comprises published scholarly papers, textbooks, proceedings of conferences, and other highly stylized documents. Our purpose is to examine the extent to which those discourse-related phenomena that generally assist the analysis of dialogue text -- where speaker, hearer, and speech-act information are more actively involved in the identification of plans and goals -- can be used to help with the analysis of expository text. In particular, we make the optimistic assumption that expository text is strongly connected; i.e., that *all* adjacent pairs of clauses in such a text are connected by 'cohesion markers,' both explicit and implicit. We investigate the impact that this assumption may have on the depth of understanding that can be achieved, the underlying semantic structures, and the supporting knowledge base for the analysis. An application of this work in designing the AI-based machine translation model, TRANSLATOR, is discussed in NIRENBURG ET AL (1986) which appears elsewhere in this volume.

When we read an expository text, our intuition relies on some basic assumptions about its coherence. That is, we normally expect the series of concepts to flow naturally from one sentence to the next. Moreover, when a conceptual discontinuity occurs at some point within the text, we are sometimes given an explicit syntactic clue (like 'on the other hand') that such will occur. More often, however, we are not given such a clue; we are expected to automatically detect this shift of focus without requiring any explicit prompting.

Most of the research in the field of discourse analysis uses texts which are dialogues; two or more people are involved, speaker and hearer roles are constantly changing, and speech-act (speaker's intention) information is a changing and essential factor in the semantics of the dialogue. For instance, extensive work has been published by LONGRACE (1977), PHILLIPS (1977), REICHMAN (1984, 1985), JOSHI ET AL. (1981), and GRIMES (1978). Although expository text does not typically contain dialogues, techniques of discourse analysis appears nevertheless to contribute strongly to the

Another area of research that directly bears upon the present problem is the notion of textual coherence. According to HOBBS (1976), an utterance is *coherent* if it is an action within the implementation of some plan. In particular, conversation may be characterized as an expression of planned behavior with goals, and is thus coherent in this sense. Hobbs describes four classes of coherent conversational moves that can occur in a dialogue: Occasion (cause or enablement), Evaluation, Explanation, and Expansion. In each of these moves, the speaker's goal is to manipulate the inference process of the hearer, so that the latter links what he/she already knows with what is new in the message. We shall illustrate that the same premise can serve as a starting point for identifying and characterizing coherence in an expository text.

2. Overview of TRANSLATOR

TRANSLATOR is the name given to an ongoing research project at Colgate University which attempts to define a basis for multilingual machine translation by using a universal intermediate metalanguage, or 'interlingua,' at its heart. The idea is to design an interlingua which is robust enough to represent sufficient syntactic, semantic, and pragmatic knowledge about a text in *any* source language, so that its translation into a different target language can proceed independently of the original text. A more thorough introduction to TRANSLATOR can be found in TUCKER AND NIRENBURG (1984) and NIRENBURG ET AL (1986).

† This material is based upon work supported by the National Science Foundation under Grant DCR-8407114.

In this paper, we limit ourselves to exploring those discourse-related phenomena which appear in expository text, and suggesting how these phenomena may be captured during the analysis of a text and represented in the interlingua itself. To support this exploration, we use those parts of the interlingua for TRANSLATOR which are relevant to discourse analysis, and identify their role in the analysis process. The use of *italics* in the paragraphs below denotes a concept which has a precise definition and connotation within interlingua itself.

An interlingua *text* may be either a single interlingua *sentence* or a series of sentences connected by *discourse operators* d. More formally:

```
text :: = sentence |
          d (text text)
```

The discourse operators d are enumerated and briefly described below; their meanings are more fully described in a later section.

Discourse Operator (d)	Use in 'd (text1 text2)'
-simil	change in topic from text1 to text2
+simil	continuation of same topic
+expan	expansion
-expan	generalization
temp	temporal sequence
condi	conditional (cause or enablement)
compare	comparison
equiv	equivalence

An interlingua *sentence* is composed of a series of *clauses*, together with its own characteristic *subworld*, *modality*, *focus*, and *speech-act* information.

Without going into further detail [see NIRENBURG ET AL (1986) for further description], we note that this representation abandons the traditional phrase-structure, dependency or other purely syntactic basis for representation, in favor of a far deeper level of representation for mechanical understanding.

3. Focus Shift in Expository Text

In expository text, the speaker and hearer roles are more or less permanently assigned to the author and the reader, respectively. The exposition is permanently under the control of the author, and the reader plays a more or less passive role throughout. Still, speech act information plays a role in this setting, in the following ways:

Definitions, as in 'Data that is stored more or less permanently in a computer we term a *database.*'

Opinions, as in 'We agree with the point of view that software piracy is illegal.'

Facts, as in 'The Symbolics LISP machine can have up to 8 megabytes of memory.'

Promises, as in 'We shall explain this subject more fully in Chapter 8.'

Advice, as in 'If you are not interested in the theoretical foundations of database management systems, you may wish to skip the next section.'

Questions, as in 'What is the tradeoff between flexibility and efficiency in comparing the relational and hierarchical database models?'

Commands, as in 'You should answer the following questions before proceeding to Chapter 2.'

Some of these speech acts are directly related to the topic under discussion, while others serve only to guide the reader through his/her planning and goal-setting activities while reading the text.

The identification of focus shift is enabled by both the underlying knowledge base and the discourse-related phenomena that appear in the text itself. At the outset of analysis, the text is viewed as a sequence of sentences, made up of clauses, each one containing a single focus, which may be either an *object* or an *event*. Both objects and events have frame-like representations and are derived from information stored in an underlying knowledge base. The knowledge base is assumed to be structured, so that relationships among specific kinds of objects and events are

revealed. These include, for instance, 'isa,' 'part-of,' 'be-agent-of,' and other links that tend to explain how primitive and compound events and objects are interrelated in the world.

A focus shift between adjacent sentences or clauses serves to signal the author's attempt to transfer the reader's attention from the given information to the new information that will be added to the presentation. The syntactic context within which such a shift might take place is arbitrary. For instance, consider the following two examples:

1. The data is shown below. Notice that some values are missing.
2. When data has missing values, it is called 'sparse'.

The first shows a shift from the focus 'data' to the focus 'missing values.' The second shows a shift from the focus 'data' to the focus 'sparse.' These illustrations show that the *kind* of shift that takes place between two adjacent foci in a text may vary. In the first sentence, the shift was one of *expansion*, while the shift in the second sentence was one of *generalization*.

From a strictly syntactic point of view, we see then that focus shift can take place regularly between adjacent clauses (sentence 2 above), adjacent sentences (sentences 1 above), and larger units of text which are adjacent. Thus, the network of focus shifts within a text may be complex.

4. Defining Discourse Cohesion Relations

The relations defined below are designed to provide a vehicle exposing the discourse structure of expository text. These relations are a variation of those developed by REICHMAN (1984) and HOBBS (1976); they differ because they are especially adapted for use in expository, rather than dialogue, types of text. The 'discourse cohesion relations' that can exist between two adjacent units of text c1 and c2 (which in turn may be clauses, sentences, or larger texts) are defined and illustrated as follows:

TEMPORAL: *temp(c1,c2)* is true if there is a temporal relationship between c1 and c2. For instance, the sentences 'It became overcast. It began to rain.' exhibit a link between the concepts of cloud cover and raining, in the sense that one happened before the other.

CONDITIONAL: *condi(c1,c2)* is true if c1 either causes or enables c2 to occur. For instance, the adjacent sentences 'It began to rain. John went indoors.' exhibit a cause-and-effect relationship between two conceptual actions, raining and going indoors.

EXPANSION: *+expan(c1,c2)* is true if c2 serves as an example or a further explanation of c1. For instance, the sentences 'The data is shown below. Notice that some values are missing.' exhibit this conceptual relationship.

GENERALIZATION: *-expan(c1,c2)* is true if c2 serves as a generalization of c1, such as in a definition. In the sentence, 'The software that allows a person to use and/or modify this data is called a DBMS,' the new concept DBMS is defined for the first time in the text, using refinements of another concept 'software' that occur through the discourse cohesion relation +expan. That is, if we identify 'software' as concept c1, 'allowing a person to use and/or modify data' as concept c2, and 'DBMS' as concept c3, then we see that the refined concept, say c1', results from +expan(c1,c2), and the new concept c3 results as from c1' through generalization; that is, -expan(c1',c3), or -expan(+expan(c1,c2),c3).

CONTRASTIVE: *-simil(c1,c2)* is true if c2 is either dissimilar or opposite from c1. For instance, consider the sentence, 'In accessing a database, the user gives English-like commands rather than Pascal-like algorithms.' Let c1 denote the concept of 'accessing a database,' c2 denote the (refined) concept of 'the user giving English-like commands,' and c3 denote the concept of 'the user giving Pascal-like algorithms.' Then we have the contrastive relation appearing in the following conceptual refinements: c1' = +expan(c1,c2) and c1'' = -expan(c1',c3). That is, c3 serves to refine the concept c1' by providing a counterexample from that which was provided in the original refinement of c1 by c2.

SIMILAR: *+simil(c1,c2)* is true if c2 is similar, but not explicitly identical, to c1. For example, consider the two sentences, ' One role of a DBMS is to provide quick access. That is, we want the user to be able to access any item in the database within a few seconds of response time.' If we let these two represent the concepts c1 and c2, respectively, we see that c2 is an approximately identical restatement of c1, and so +simil(c1,c2) is true.

EQUIVALENT: *equiv(c1,c2)* is true if we can further ascertain that c2 is equivalent, or conceptually identical, to c1. Often this equivalence is marked by an explicit sign of synonymy, such as the parentheses in the following example. 'The software that allows the user to access this data is called a database management system (DBMS).' Here, equivalence is marked between the newly-defined concept 'database management system' and the acronym DBMS.

DIGRESSION: *none(c1,c2)* is true if none of the other relations listed above exist between c1 and c2.

5. Inferring Focus Shift and Discourse Relations

Following the definition of these discourse cohesion classes, it is necessary to identify some principles upon which the discourse structure may be revealed in the text as analysis progresses from the first sentence forward. That is, at any point in the reading of a text, the system must understand 'what's going on' in the sense of its discourse structure.

Letting c1 and c2 again denote a pair of items which appear adjacent to each other in a text, the following principles can be used to identify focus shift, based on the discourse cohesion relations that can occur between c1 and c2.

1. If c1 is followed by c2 and +expan(c1,c2) is true, then a focus shift from c1 to c1' takes place. That is, c1' is an embellishment of c1 due to the relationship +expan and the supporting concept c2.
2. Similarly, the relation -simil(c1,c2) yields the focus shift from c1 to the embellishment c1'.
3. If c1 is followed by c2 and -expan(c1,c2) is true, then the focus shift from c1 to c2 takes place. That is, c1 relinquishes its role as the focus of discourse to c2 by the process of generalization.
4. Similarly, each one of the relations condi(c1,c2), temp(c1,c2), and none(c1,c2) yields a focus shift from c1 to c2.
5. On the other hand, the relations +simil(c1,c2) and equiv(c1,c2) cause no shift to take place; that is, c1 remains the focus of discourse after c2 has been processed in each case.

Connectivity between adjacent concepts in a text is sometimes explicitly revealed by the presence of 'clue words' and other markers. The use of clue words for discourse analysis is common (eg REICHMAN (1984)). The example text discussed in the following section contains several such clue words. Sometimes the marker appears as a punctuation mark (such as a parenthetical which signals the relation +equiv), other instances appear as single words (such as 'However' signaling -simil), while still others are complete clauses (such as 'there may be far less' signaling +simil).

Yet, many instances of conceptual connectivity are *not* cued by the presence of such markers; the are revealed instead by general syntactic structure (such as the appearance of a relative clause, signaling +expan) or by semantic properties that are possessed by the underlying concepts and stored in the knowledge base. The following discussion suggests how such knowledge can be used to mark instances of conceptual connectivity in expository text.

Intuitively, some of the conceptual properties that reveal discourse cohesion relations are the following:

Property	Connective
isa	-expan
agent, agent-of	+expan
object, object-of	+expan
patient, patient-of	+expan
instrument, instrument-of	+expan
source, source-of	+expan
destination, destination-of	+expan
time	temp
space	+expan
effects	condi

Merging these conceptual clues with the explicit syntactic clues for discourse connectivity, leads to the following table. This table reveals some of the clues (both explicit and implicit) that lead to exposure of the

cohesion relation d(c1 c2), where c1 and c2 are adjacent concepts (processes or objects) within the text.

Syntactic clues (explicit)	Conceptual clues (implicit)	Relation d(c1 c2)
c1 'then' c2	time(c1) precedes time(c2)	temp(c1,c2)
'if' c1 'then' c2 or c1 'caused' c2 or c1 'enabled' c2	c2 in effects(c1)	condi(c1,c2)
c2 in relative clause for c1	c2 in properties(c1) or c2 in links(c1) or c2 isa(c1)	+expan(c1,c2)
c1 'is' c2	c1 is-part-of(c2)	-expan(c1,c2)
c1. 'However,' c2		-simil(c1,c2)
c1 'is like' c2		+simil(c1,c2)
c1 (c2)	c1 = c2	equiv(c1,c2)

A simple algorithm to infer such relations between pairs of concepts in the text, ci and cj, can be given. However, space does not permit its further elaboration in this paper.

6. An Example

To illustrate the application of these ideas, we have analyzed the five sentences of a paragraph taken from the first page of Jeffrey Ullman's book, *Principles of Database Systems*, given below in a specially annotated form. The annotations C, S, and D on the left denote clauses, sentences, and discourse cohesion markers that are uncovered in a parse of this paragraph.

Identification			Concept or Connective
	S1		
		C1	Data,
D			such as the above,
		C2	that is stored more-or-less permanently in a computer
D			we term
		C3	a database.
	S2		
		C4	The software
		C5	T{that allows one or many persons to use or modify this data
D			is
		C6	a database management system
D			( ).
		C7	DBMS
	S3		
		C8	A major role of the DBMS
D			is
		C9	to allow the user to deal with the data
		C10	in abstract terms,
D			rather than
		C11	... [to allow the user to deal with the data]
		C12	as the computer stores the data.
	S4		
D			In this sense,
		C13	the DBMS
D			acts as
		C14	an interpreter for a high-level programming language,
		C15	ideally allowing the user to specify what must be done,
D			with little or no attention on the user's part
		C16	to the detailed algorithms or data representation
D			used by
		C17	the system.
	S5		
D			However,
		C18	in the case of a DBMS,
D			there may be far less
		C19	relationship between the data as seen by the user and ...[the data] as stored by the computer
D			than
		C20	...[the relationship] between, say, arrays as defined in a typical programming language and the representation of those arrays in memory.

While space does not permit a detailed description of the analysis of this text, below is a summarization of the final result of such an analysis.

	New Focus	Derived From	Derived Concept (in CAPS)
S1			DATABASE
	C1'	+expan(C1,C2)	DATA
D			such as the above,
	C3	-expan(C1',C3)	DATABASE
S2			DATABASE SYSTEM
	C6	-expan(C4',C6)	DATABASE SYSTEM
S3			ROLE OF DBMS
	C8''	-simil(C8',C11')	ROLE OF DBMS
S4			ROLE OF DBMS
D			In this sense,
	C13'''	-simil(C13'',C16')	ROLE OF DBMS
S5			RELATIONSHIP OF DATA
D			However,
	C19''	-simil(C19',C20)	RELATIONSHIP OF DATA

Here, we note that each sentence has inherited a focus, and the remaining connectives and semantic properties can later be used to expose the overall discourse structure of the paragraph.

7. Conclusion

We have outlined a basis for modeling semantic connectivity among clauses and sentences in an expository text. Strong notions of discourse relations, focus, and an underlying knowledge base are essential to this process.

REFERENCES

1. Grimes, J., 'Topic Levels,' *Theoretical Issues in Natural Language Processing 2*, Association for Computational Linguistics, 1978.

2. Hobbs, J., 'A Computational Approach to Discourse Analysis,' City University of New York (1976).

3. Joshi, A., B. Webber and I. Sag (eds), *Elements of Discourse Understanding*, Cambridge University Press (1981).

4. Longrace,R. and S. Levinsohn, 'Field Analysis of Discourse,' in W. Dressler (ed), *Current Trends in Text Linguistics*, DeGruyter (1977).

5. Nirenburg, Sergei, Victor Raskin, and Allen Tucker, 'Interlingua Design for TRANSLATOR,' *Proceedings of the Conference on Theoretical and Methodological Issues in Machine Translation*, Colgate University (August 14-16, 1985), 224-244.

6. Nirenburg, Sergei, Victor Raskin and Allen Tucker, 'On Knowledge Based Machine Translation,' *Proceedings of COLING 86*.

7. Phillips, B., 'Discourse Connectives,' Technical Report KS-11, Department of Engineering, University of Illinois at Chicago, 1977.

8. Raskin, Victor and Sergei Nirenburg, 'A Metric for Computational Analysis of Meaning,' *Proceedings of COLING 86*.

9. Reichman, Rachel, 'Extended Person-Machine Interface, *Artificial Intelligence 22* (1984), 157-218.

10. Tucker, Allen and Sergei Nirenburg, 'Machine Translation: A Contemporary View,' in M. Williams (ed), *Annual Review of Information Science and Technology 19*, Knowledge Industry Publications (October, 1984), 129-160.