

Semantics of Paragraphs

Wlodek Zadrozny*

Karen Jensen*

IBM T. J. Watson Research Center

We present a computational theory of the paragraph. Within it we formally define coherence, give semantics to the adversative conjunction “but” and to the Gricean maxim of quantity, and present some new methods for anaphora resolution.

The theory precisely characterizes the relationship between the content of the paragraph and background knowledge needed for its understanding. This is achieved by introducing a new type of logical theory consisting of an object level, corresponding to the content of the paragraph, a referential level, which is a new logical level encoding background knowledge, and a metalevel containing constraints on models of discourse (e.g. a formal version of Gricean maxims). We propose also specific mechanisms of interaction between these levels, resembling both classical provability and abduction. Paragraphs are then represented by a class of structures called p-models.

1. Introduction

Logic and knowledge have been often discussed by linguists. Anaphora is another prominent subject in linguistic analyses. Not so frequently examined are different types of cohesion. And it is quite rare to find the word “paragraph” in articles or books about natural language understanding, although paragraphs are grammatical units and units of discourse. But it is possible to speak formally about the role of background knowledge, cohesion, coherence and anaphora—all within one, flexible and natural, logical system—if one examines the semantic role of the linguistic construct called a paragraph.

Paragraphs have been sometimes described, rather loosely, as “units of thought.” We establish a correspondence between them and certain types of logical models, thereby making the characterization of paragraphs more precise. The correspondence gives us also an opportunity to identify and attack—with some success, we believe—three interesting and important problems: (1) how to define formally *coherence* and *topic*, (2) how to resolve *anaphora*, and (3) what is the formal meaning of *linkages* (*connectives*) such as *but*, *however*, *and*, *certainly*, *usually*, *because*, *then*, etc. These questions are central from our point of view because: (1) the “unity” of a paragraph stems from its coherence, while the “aboutness” of thought can be, at least to some extent, described as existence of a topic; (2) without determining reference of pronouns and phrases, the universes of the models are undefined; and (3) the linkages, which make sentences into paragraphs, have semantic roles that must be accounted for. We can explain then the process of building a computational model of a paragraph (a *p-model*) as an interaction between its sentences, background knowledge to which these sentences refer, and metatheoretical operators that indicate types of permitted models.

* P.O. Box 704, Yorktown Heights, NY 10598

At this point the reader may ask: what is so special about paragraphs; does all this mean that a chapter, a book or a letter do not have any formal counterparts? We believe they do. But we simply do not yet know how corresponding formal structures would be created from models of paragraphs. To answer this question we may need more advanced theories of categorization and learning than exist today. On the other hand, the paragraph is the right place to begin: it is the next grammatical unit after the sentence; connectives providing cohesion operate here, not at the level of an entire discourse; and it is the smallest reasonable domain of anaphora resolution. Furthermore, we will argue, it is the smallest domain in which topic and coherence can be defined.

The formalization of paragraph structure requires the introduction of a new type of logical theory and a corresponding class of models. As we know, the usual logical structures consist of an *object level* theory \mathbf{T} and provability relation \vdash ; within the context of the semantics of natural language, the object theory contains a logical translation of the surface form of sentences, and \vdash is the standard provability relation (logical consequence). In mathematical logic, this scheme is sometimes extended by adding a metalevel assumption, for instance postulating the standardness of natural numbers. In artificial intelligence, a metarule—typically, the closed world assumption of circumscription—can be used in dealing with theoretical questions, like the frame problem. But a formal account of natural language understanding requires more. It requires at least a description (a) of how background knowledge about objects and relations that the sentences describe is used in the process of understanding, and (b) of general constraints on linguistic communications, as expressed for instance in Gricean maxims. It is well known that without the former it is impossible to find references of pronouns or attachments of prepositional phrases; background knowledge, as it turns out, is also indispensable in establishing coherence. We have then reasons for introducing a new logical level—a *referential level* \mathbf{R} , which codes the background knowledge. As for Gricean maxims, we show that they can be expressed formally and can be used in a computational model of communication. We include them in a *metalevel* \mathbf{M} , which contains global constraints on models of a text and definitions of meta-operators such as the conjunction *but*. We end up with three-level logical theories $(\mathbf{M}, \mathbf{T}, \mathbf{R}, \vdash_{\mathbf{R} + \mathbf{M}})$, where a provability relation $\vdash_{\mathbf{R} + \mathbf{M}}$, based on \mathbf{R} and \mathbf{M} , can be used, for example, to establish the reference of pronouns.

This work is addressed primarily to our colleagues working on computational models of natural language; but it should be also of interest to linguists, logicians, and philosophers. It should be of interest to linguists because the notion that a paragraph is equivalent to a model is something concrete to discuss; because p-models are as formal as formal languages (and therefore something satisfyingly theoretical to argue about); and because new directions for analysis are opened beyond the sentence. The work should be of interest to logicians because it introduces a new type of three-level theory, and corresponding models. The theory of these structures, which are based on linguistic constructs, will differ from classical model theory—for instance, by the fact that names of predicates of an object theory matter, because they connect the object theory with the referential level. This work should be of interest to philosophers for many of the same reasons: it makes more sense to talk about the meaning of a paragraph than about the meaning of a sentence. The following parallel can be drawn: a sentence is meaningful only with respect to a model of a paragraph, exactly as the truth value of a formula can be computed only with respect to a given model. Moreover, it is possible in this framework to talk about meaning without mentioning the idea of possible worlds. However, we do not identify meaning with truth conditions; in this paper, the meaning of a sentence is its role in the model of the paragraph

in which this sentence occurs. Our intuitive concept of meaning is similar to Lakoff's (1987) Idealized Cognitive Model (ICM). Needless to say, we believe in the possibility of formalizing ICMs, although in this paper we will not try to express, in logic, prototype effects, metaphors, or formal links with vision.

The paper is presented in six sections. In Section 2, we discuss the grammatical function of the paragraph and we show, informally, how a formal model of a paragraph might actually be built. In Section 3 we give the logical preliminaries to our analysis. We discuss a three-part logical structure that includes a referential level, and we introduce a model for plausible meaning. Section 4 discusses paragraph coherence, and Section 5 constructs a model of a paragraph, a *p-model*, based on the information contained in the paragraph itself and background information contained in the referential level *R*. Section 5 further motivates the use of the referential level, showing how it contributes to the resolution of anaphoric reference. In Section 6, we broaden our presentation of the metalevel, introducing some metalevel axioms, and sketching ways by which they can be used to reduce ambiguity and construct new models. We also show metalevel rules for interpreting "but."

2. The Paragraph as a Discourse Unit

2.1 Approaches to Paragraph Analysis

Recent syntactic theory—that is, in the last 30 years—has been preoccupied with sentence-level analysis. Within discourse theory, however, some significant work has been done on the analysis of written paragraphs. We can identify four different linguistic approaches to paragraphs: prescriptivist, psycholinguist, textualist, and discourse-oriented.

The prescriptivist approach is typified in standard English grammar textbooks, such as Warriner (1963). In these sources, a paragraph is notionally defined as something like a series of sentences that develop one single topic, and rules are laid down for the construction of an ideal (or at least an acceptable) paragraph. Although these dictates are fairly clear, the underlying notion of topic is not.

An example of psycholinguistically oriented research work can be found in Bond and Hayes (1983). These authors take the position that a paragraph is a psychologically real unit of discourse, and, in fact, a formal grammatical unit. Bond and Hayes found three major formal devices that are used, by readers, to identify a paragraph: (1) the repetition of content words (nouns, verbs, adjectives, adverbs); (2) pronoun reference; and (3) paragraph length, as determined by spatial and/or sentence-count information. Other psycholinguistic studies that confirm the validity of paragraph units can be found in Black and Bower (1979) and Haberlandt et al. (1980).

The textualist approach to paragraph analysis is exemplified by E. J. Crothers. His work is taxonomic, in that he performs detailed descriptive analyses of paragraphs. He lists, classifies, and discusses various types of inference, by which he means, generally, "the linguistic-logical notions of consequent and presupposition" Crothers (1979:112) have collected convincing evidence of the existence of language chunks—real structures, not just orthographic conventions—that are smaller than a discourse, larger than a sentence, generally composed of sentences, and recursive in nature (like sentences). These chunks are sometimes called "episodes," and sometimes "paragraphs." According to Hinds (1979), paragraphs are made up of segments, which in turn are made up of sentences or clauses, which in turn are made up of phrases. Paragraphs therefore give hierarchical structure to sentences. Hinds discusses three major types of paragraphs, and their corresponding segment types. The three types are procedural (how-to), expository (essay), and narrative (in this case, spontaneous conversation). For each type,

its segments are distinguished by bearing distinct relationships to the paragraph *topic* (which is central, but nowhere clearly defined). Segments themselves are composed of clauses and regulated by “switching” patterns, such as the question-answer pattern and the remark-reply pattern.

2.2 Our View of Paragraphs: An Informal Sketch

Although there are other discussions of the paragraph as a central element of discourse (e.g. Chafe 1979, Halliday and Hasan 1976, Longacre 1979, Haberlandt et al. 1980), all of them share a certain limitation in their formal techniques for analyzing paragraph structure. Discourse linguists show little interest in making the structural descriptions precise enough so that a computational grammar of text could adapt them and use them. Our interest, however, lies precisely in that area.

We suggest that the paragraph is a grammatical and logical unit. It is the smallest linguistic representation of what, in logic, is called a “model,” and it is the first reasonable domain of anaphora resolution, and of coherent thought about a central topic.

A paragraph can be thought of as a grammatical unit in the following sense: it is the discourse unit in which a functional (or a predicate-argument) structure can be definitely assigned to sentences/strings. For instance, Sells (1985, p. 8) says that the sentence “Reagan thinks bananas,” which is otherwise strange, is in fact acceptable if it occurs as an answer to the question “What is Kissinger’s favorite fruit?” The pairing of these two sentences may be said to create a small paragraph. Our point is that an acceptable structure can be assigned to the utterance “Reagan thinks bananas” only within the paragraph in which this utterance occurs. We believe that, in general, no unit larger than a paragraph is necessary to assign a functional structure to a sentence, and that no smaller discourse fragment, such as two (or one) neighboring sentences, will be sufficient for this task. That is, we can ask in the first sentence of a paragraph about Kissinger’s favorite fruit, elaborate the question and the circumstances in the next few sentences, and give the above answer at the end. We do not claim that a paragraph is necessarily described by a set of grammar rules in some grammar formalism (although it may be); rather, it has the grammatical role of providing functional structures that can be assigned to strings.

The logical structure of paragraphs will be analyzed in the next sections. At this point we would like to present some intuitions that led to this analysis. But first we want to identify our point of departure. In order to resolve anaphora and to establish the coherence or incoherence of a text, one must appeal to the necessary background knowledge. Hence, a formal analysis of paragraphs must include a formal description of background knowledge and its usage. Furthermore, this background knowledge cannot be treated as a collection of facts or formulas in some formal language, because that would preclude dealing with contradictory word senses, or multiple meanings. Secondly, this background knowledge is not infinite and esoteric. In fact, to a large extent it can be found in standard reference works such as dictionaries and encyclopedias.

To argue for these points, we can consider the following paragraph:¹

In the summer of 1347 a merchant ship returning from the Black Sea entered the Sicilian port of Messina bringing with it the horrifying disease that came to be known as the Black Death. It struck rapidly. Within twenty-four hours of infection and the appearance of the first small black pustule came an agonizing death. The effect of the Black Death was appalling. In less than twenty years half

1 J. Burke, *The Day the Universe Changed*. 1986. Little, Brown & Co., Boston, Massachusetts, p. 55.

the population of Europe had been killed, the countryside devastated, and a period of optimism and growing economic welfare had been brought to a sudden and catastrophic end.

The sentences that compose a paragraph must stick together; to put it more technically, they must cohere. This means very often that they show cohesion in the sense of Halliday (1976)—semantic links between elements. Crucially, also, the sentences of a paragraph must all be related to a topic.

However, in the example paragraph, very few instances can be found here of the formal grammatical devices for paragraph cohesion. There are no connectives, and there are only two anaphoric pronouns (both occurrences of "it"). In each case, there are multiple possible referents for the pronoun. The paragraph is coherent because it has a topic: "Black Death"; all sentences mention it, explicitly or implicitly.

Notice that resolving anaphora precedes the discovery of a topic. A few words about this will illustrate the usage of background knowledge. By parsing with syntactic information alone, we show that resolution of the first "it" reference hinges on the proper attachment of the participial clause "bringing within it...". If the "bringing" clause modifies "Messina," then "Messina" is the subject of "bringing" and must be the referent for "it." If the clause modifies "port," then "port" is the desired referent; if the clause is attached at the level of the main verb of the sentence, then "ship" is the referent.

But syntactic relations do not suffice to resolve anaphora: Hobbs' (1976) algorithm for resolving the reference of pronouns, depending only on the surface syntax of sentences in the text, when applied to "it" in the example paragraph, fails in both cases to identify the most likely referent NP.

Adding selectional restrictions (semantic feature information, Hobbs 1977) does not solve the problem, because isolated features offer only part of the background knowledge necessary for reference disambiguation. Later, Hobbs (1979, 1982) proposed a knowledge base in which information about language and the world would be encoded, and he emphasized the need for using "salience" in choosing facts from this knowledge base.

We will investigate the possibility that the structure of this knowledge base can actually resemble the structure of, for example, natural language dictionaries. The process of finding referents could then be automated.

Determining that the most likely subject for "bringing," in the first sentence, is the noun "ship" is done in the following fashion. The first definition for "bring" in W7 (*Webster's Seventh Collegiate Dictionary*) is "to convey, lead, carry, or cause to come along with one..." The available possible subjects for "bringing" are "Messina," "port," and "ship." "Messina" is listed in the *Pronouncing Gazetteer* of W7, which means that it is a place (and is so identified in the subtitle of the *Gazetteer*). So we can substitute the word "place" for the word "Messina." Then we check the given definitions for the words "place," "port," and "ship" in both dictionaries. LDOCE (*Longman Dictionary of Contemporary English*) proves particularly useful at this point. Definitions for "place" begin: "a particular part of space...". Definitions for "port" include: "harbour..."; "an opening in the side of a ship...". But the first entry for "ship" in LDOCE reads "a large boat for carrying people or goods...". This demonstrates a very quick connection with the definition for the verb "bring," since the word "carry" occurs in both definitions. It requires much more time and effort to find a connection between "bring" and either of the other two candidate subject words "place" or "port." Similar techniques can be used to assign "disease" as the most probable referent for the second "it" anaphor in our example paragraph.

Equally significant in this instance is the realization that a dictionary points to synonym and paraphrase relations, and thereby verifies the cohesiveness of the passage. Through the dictionary (LDOCE again), we establish shared-word relationships between and among the words “disease,” “Black death,” “infection,” “death,” “killed,” and “end.” Note that there is no other means, short of appealing to human understanding or to some hand-coded body of predicate assertions, for making these relationships.

This demonstrates that information needed to identify and resolve anaphoric references can be found, to an interesting extent, in standard dictionaries and thesauri. (Other reference works could be treated as additional sources of world knowledge.) This type of consultation uses existing natural language texts as a *referential level* for processing purposes. It is the lack of exactly this notion of referential level that has stood in the way of other linguists who have been interested in the paragraph as a unit. Crothers (1979, p. 112), for example, bemoans the fact that his “theory lacks a world knowledge component, a mental ‘encyclopedia,’ which could be invoked to generate inferences...”. With respect to that independent source of knowledge, our main contributions are two. First, we identify its possible structure (a collection of partially ordered theories) and make formal the choice of a most plausible interpretation. In other words, we recognize it as a separate logical level—the referential level. Second, we suggest that natural language reference works, like dictionaries and thesauri, can quite often fill the role of the referential level.

3. The Logic of Reference

The goal of this section is to introduce a formalism describing how background knowledge is used in understanding text. The term “logic of reference” denotes a formal description of this process of consulting various sources of information in order to produce an interpretation of a text. The formalist will be presented in a number of steps in which we will elaborate one simple example:

Example 1

Entering the port, a ship brought a disease.

This sentence can be translated into the logical formula (ignoring only the past tense of “bring” and the progressive of “enter”):

Definition S: $enter(x1, x2) \ \& \ ship(x1) \ \& \ port(x2) \ \& \ bring(x3, x4) \ \& \ disease(x4) \ \& \ x1 = s$
 $\ \& \ x2 = m \ \& \ x3 = s \ \& \ x4 = d,$

where $s, m, d,$ are constants.

We adopt the three-level semantics as a formal tool for the analysis of paragraphs. This semantics was constructed (Zadrozny 1987a, 1987b) as a formal framework for default and commonsense reasoning. It should not come as a surprise that we can now use this apparatus for text/discourse analysis; after all, many natural language inferences are based on defaults, and quite often they can be reduced to choosing most plausible interpretations of predicates. For instance, relating “they” to “apples” in the sentence (cf. Haugeland 1985 p. 195; Zadrozny 1987a):

We bought the boys apples because they were so cheap

can be an example of such a most plausible choice.

The main ideas of the three-level semantics can be stated as follows:

1. Reasoning takes place in a three-level structure consisting of an *object level*, a *referential level*, and a *metalevel*.
2. The *object level* is used to describe the current situation, and in our case is reserved for the formal representation of paragraph sentences. For the sake of simplicity, the object level will consist of a first order theory.
3. The *referential level*, denoted by **R**, consists of theories representing background knowledge, from which information relevant to the understanding of a given piece of text has to be extracted. It constrains interpretations of the predicates of an object theory. Its structure and the extraction methods will be discussed below.
4. Understanding has as its goal construction of an interpretation of the text, i.e. building some kind of model. Since not all logically permissible models are linguistically appropriate, one needs a place, namely the *metalevel*, to put constraints on types of models. Gricean maxims belong there; Section 6 will be devoted to a presentation of the metalevel rules corresponding to them.

We have shown elsewhere (Jensen and Binot 1988; Zadrozny 1987a, 1987b) that natural language programs, such as on-line grammars and dictionaries, can be used as referential levels for commonsense reasoning—for example, to disambiguate PP attachment. This means that information contained in grammars and dictionaries can be used to constrain possible interpretations of the *logical* predicates of an object-level theory.

The referential structures we are going to use are collections of logical theories, but the concept of reference is more general. Some of the intuitions we associate with this notion have been very well expressed by Turner (1987, pp. 7–8):

...Semantics is constrained by our models of ourselves and our worlds. We have models of *up* and *down* that are based by the way our bodies actually function. Once the word “up” is given its meaning relative to our experience with gravity, it is not free to “slip” into its opposite. “Up” means up and not down. ... We have a model that men and women couple to produce offspring who are similar to their parents, and this model is grounded in genetics, and the semantics of kinship metaphor is grounded in this model. Mothers have a different role than fathers in this model, and thus there is a reason why “Death is the father of beauty” fails poetically while “Death is the mother of beauty” succeeds....

It is precisely this “grounding” of logical predicates in other conceptual structures that we would like to capture. We investigate here only the “grounding” in logical theories. However, it is possible to think about constraining linguistic or logical predicates by simulating physical experiences (cf. Woods 1987).

We assume here that a translation of the surface forms of sentences into a logical formalism is possible. Its details are not important for our aim of giving a semantic interpretation of paragraphs; the main theses of our theory do not depend on a logical notation. So we will use a very simple formalism, like the one above, resembling the standard first order language. But, obviously, there are other possibilities—for instance, the discourse representation structures (DRS's) of Kamp (1981), which have been used to translate a subset of English into logical formulas, to model text (identified with a list of sentences), to analyze a fragment of English, and to deal with anaphora. The logical

notation of Montague (1970) is more sophisticated, and may be considered another possibility. Jackendoff's (1983) formalism is richer and resembles more closely an English grammar. Jackendoff (1983, p. 14) writes "it would be perverse not to take as a working assumption that language is a relatively efficient and accurate encoding of the information it conveys." Therefore a formalism of the kind he advocates would probably be most suitable for an implementation of our semantics. It will also be a model for our simplified logical notation (cf. Section 5). We can envision a system that uses data structures produced by a computational grammar to obtain the logical form of sentences.

3.1 Finite Representations, Finite Theories

Unless explicitly stated otherwise, we assume that formulas are expressed in a certain (formal) language L **without** equality; the extension $L(=)$ of L is going to be used only in Section 5 for dealing with noun phrase references. This means that natural language expressions such as "A is B," "A is the same as B," etc. are not directly represented by logical equality; similarly, "not" is often not treated as logical negation; cf. Hintikka (1985).

All logical notions that we are going to consider, such as theory or model, will be finitary. For example, a model would typically contain fewer than a hundred elements of different logical sorts. Therefore these notions, and all other constructs we are going to define (axioms, metarules, definitions etc.) are computational, although usually we will not provide explicit algorithms for computing them. The issues of control are not so important for us at this point; we restrict ourselves to describing the logic. This *Principle of Finitism* is also assumed by Johnson-Laird (1983), Jackendoff (1983), Kamp (1981), and implicitly or explicitly by almost all researchers in computational linguistics. As a logical postulate it is **not** very radical; it is possible within a finitary framework to develop that part of mathematics that is used or has potential applications in natural science, such as mathematical analysis (cf. Mycielski 1981).

On the other hand, a possible obstacle to our strategy of using only finite objects is the fact that the deductive closure of any set of formulas is not finite in standard logic, while, clearly, we will have to deduce new facts from formal representations of text and background knowledge. But there are several ways to avoid this obstruction. For example, consider theories consisting of universal formulas without function symbols. Let $Th(T)$ of such a theory T be defined as T plus ground clauses/sets of literals provable from T in standard logic. It is easily seen that it is a closure, i.e. $Th(Th(T)) = Th(T)$; and obviously, it is finite, for finite T . It makes sense then to require that logical consequences of paragraph sentences have similar finite representations. However, in order not to limit the expressive power of the formal language, we should proceed in a slightly different manner. The easiest way to achieve the above requirement is by postulating that all universes of discourse are always finite, and therefore all quantifiers actually range over finite domains. In practice, we would use those two and other tricks: we could forbid more than three quantifier changes, because even in mathematics more than three are rare; we could restrict the size of universes of discourse to some large number such as 1001; we could allow only a fixed finite nesting of function symbols (or operators) in formulas; etc. The intention of this illustration was to convince the reader that we now can introduce the following set of definitions.

Definitions

- A *theory* is a finite set of sentences *Sent* (formulas without free variables in some formal language).

- A *deductive closure* operator is a function $Th : \mathbf{P}(\text{Sent}) \rightarrow \mathbf{P}(\text{Sent})$
 - (a) $T \subset Th(T)$, for any T ,
 - (b) $Th(Th(T)) = Th(T)$,
 - (c) $Th(T)$ is finite, for finite T ; additionally, we require it to be ground, for ground T .
- A theory T is *consistent* if there is no formula ϕ such that both ϕ and $\neg\phi$ belong to $Th(T)$ (and inconsistent otherwise).
- A *model* of a theory T is defined, as usual, as an interpretation defined on a certain domain, which satisfies all formulas of T . The collection of all (finite) models of a theory T will be denoted by $Mods(T)$.
- The set of all subformulas of a collection of formulas F is denoted by $Form(F)$. $\bar{\phi}$ is a *ground instance* of a formula ϕ , if $\bar{\phi}$ contains no variables, and $\bar{\phi} = \phi\theta$, for some substitution θ .

Thus, we do not require $Th(T)$ to be closed under substitution instances of tautologies. Although in this paper we take *modus ponens* as the main rule of inference, in general one can consider deductive closures with respect to weaker, nonstandard logics, (cf. Levesque 1984; Frisch 1987; Patel-Schneider 1985). But we won't pursue this topic further here.

3.2 The Structure of Background Knowledge

Background knowledge is not a simple list of meaning postulates—it has a structure and it may contain contradictions and ambiguities. These actualities have to be taken into account in any realistic model of natural language understanding. For instance, the verb “enter” is polysemous. But, unless context specifies otherwise, “to come in” is a more plausible meaning than “to join a group.” Assuming some logical representation of this knowledge, we can write that

$$enter(x, y) \rightarrow \{come_in(x, y); place(y)\} \quad (e1)$$

$$enter(x, y) \rightarrow \{join(x, y)\&group(y); \dots\} \quad (e2)$$

and $e2 <_{enter} e1$.

Two things should be explained now about this notation:

- Meanings of predicates/words are represented on the right-hand sides of the arrows as collections of formulas—i.e., theories. The main idea is that these mini-theories of predicates appearing in a paragraph will jointly provide enough constraints to exclude implausible interpretations. (One can think of meanings as regions in space, and of constraints as sets of linear inequalities approximating these regions). How this can be done, we will show in a moment.
- These theories are partially ordered; and their partial orders are written as $<_{enter(x,y)}$, or $<_{enter}$, or $<_i$, or simply $<$, depending on context. This is our way of making formal the asymmetry of plausibility of different meanings of a predicate. Again, a way of exploiting it will be shown below.

Definition

- A *referential level* \mathbf{R} is a structure

$$\mathbf{R} = \{(\psi, <_{\psi}) : \psi \in \text{Formulae}\}$$

where—for each ψ — $<_{\psi}$ is a partially ordered (by a relation of plausibility) collection of implications $\psi \rightarrow T_{\psi}$.

The term $\psi \rightarrow T$ stands for the theory $\{\psi \rightarrow \tau : \tau \in T\}$, and $\psi \rightarrow \{\phi_1, \phi_2, \dots\}$ abbreviates $\{\psi \rightarrow \phi_1, \psi \rightarrow \phi_2, \dots\}$.

- It is convenient to assume also that all formulas, except “laws”—which are supposed to be always true—have the least preferred *empty interpretation* \tilde{a} .
- We suppose also that interpretations are additionally ranked according to the canonical partial ordering on subformulas. The ranking provides a natural method of dealing with exceptions, as in the case of finding an interpretation of $\alpha \ \& \ \rho \ \& \ \beta$ with \mathbf{R} containing $(\alpha \rightarrow y)$, $(\alpha \ \& \ \beta \rightarrow \neg y)$, where $\neg y$ would be preferred to y —if both are consistent, and both defaults are equally preferred. This means that preference is given to more specific information. For instance, the sentence *The officer went out and struck the flag* will get the reading “lowered the flag,” if the appropriate theory of $\text{strike}(x, y) \ \& \ \text{flag}(y)$ is part of background knowledge; if not, it will be understood as “hit the flag.”

The referential level $(\mathbf{R}, <)$ may contain the theories listed below. Since we view a dictionary as an (imperfect) embodiment of a referential level, we have derived the formulas in every theory T_{ψ} from a dictionary definition of the term ψ . We believe that even such a crude model can be useful in practice, but a refinement of this model will be needed to have a sophisticated theory of a working natural language understanding system.

$$\text{enter}(x, y) \rightarrow \{\text{come_in}(x, y); \text{place}(y); \dots\} \quad (\text{e1})$$

/* enter—to come into a place */

$$\text{enter}(x, y) \rightarrow \{\text{join}(x, y) \ \& \ \text{group}(y); \text{typically} : \text{professionals}(y)\} \quad (\text{e2})$$

/* enter—to join a group; typically of professionals */

$$\text{ship}(x) \rightarrow \{\text{large_boat}(x); \exists y \text{carry}(x, y) \ \& \ (\text{people}(y) \vee \text{goods}(y)); \dots\} \quad (\text{sh1})$$

/* ship—a large boat for carrying people or goods on the sea */

$$\text{ship}(x) \rightarrow \{(\text{large_aircraft}(x) \vee \text{space_vehicle}(x)); \dots\} \quad (\text{sh2})$$

$$\text{bring}(x, y) \rightarrow \{\text{carry}(x, y); \dots\} \quad (\text{b1})$$

/* bring—to carry */

$$\text{bring}(x, y) \rightarrow \{\text{cause}(x, y); \dots\} \quad (\text{b2})$$

/* bring—to cause */

$$\text{disease}(y) \rightarrow \{\text{illness}(y); \dots\} \quad (\text{de1})$$

/* disease—illness caused by an infection */

$$\begin{aligned}
 \text{disaster}(y) &\rightarrow \{ \dots; \exists x \text{ cause}(y, x) \& \text{harm}(x) \dots \} && (\text{dr1}) \\
 &&& /* \text{disaster} \text{---} \text{causes a harm} */ \\
 \text{port}(x) &\rightarrow \{ \text{harbor}(x); \dots \} && (\text{p1}) \\
 &&& /* \text{port} \text{---} \text{harbor} */ \\
 (\dots) &&&
 \end{aligned}$$

Note: We leave undefined the semantics of adverbs such as **typically** in (e2). This adverb appears in the formula as an operator; our purpose in choosing this representation is to call the reader's attention to the fact that for any real applications the theories will have to be more complex and very likely written in a higher order language (cf. Section 4).

The theories, which we describe here only partially, restricting ourselves to their relevant parts, represent the meanings of concepts. We assume as before that (e1) is more plausible than (e2), i.e. $e2 <_{\text{enter}} e1$; similarly, for (sh1), (sh2) and $<_{\text{ship}}$, etc. This particular ordering of theories is based on the ordering of meanings of the corresponding words in dictionaries (derived and less frequent meanings have lower priority). But one can imagine deriving such orderings by other means, such as statistics.

The partial order $<_{\text{enter}}$ has the theories $\{e1, e2, \tilde{a}\}$ as its domain; \tilde{a} is the least preferred empty interpretation corresponding to our lack of knowledge about the predicate; it is used when both (e1) and (e2) are inconsistent with a current object theory. The domain is ordered by the relation of preference $\tilde{a} <_{\text{enter}} e2 <_{\text{enter}} e1$. The theory (e1) will always be used in constructing theories and models of paragraphs in which the expression "enter" (in any grammatical form) appears, unless assuming it would lead to an inconsistency. In such a case, the meaning of "to enter" would change to (e2), or some other theory belonging to **R**.

We would like to stress three points now: (1) the above implications are based on the definitions that actually occur in dictionaries; (2) the ordering can actually be found in some dictionaries—it is not our own arbitrary invention; (3) it is natural to treat a dictionary definition as a **theory**, since it expresses "the analysis of a set of facts in their relation to one another," different definitions corresponding to possible different analyses. (Encyclopedia articles are even more theory-like.) In this sense, the notion of a referential level is a formalization of a real phenomenon.

Obviously, dictionaries or encyclopedias do not include all knowledge an agent must have to function in the world, or a program should possess in order to understand any kind of discourse. Although the exact boundary between world knowledge and lexical knowledge is impossible to draw, we do know that lexicons usually contain very little information about human behavior or temporal relations among objects of the world. Despite all these differences, we may assume that world knowledge and lexical knowledge (its proper subset) have a similar formal structure. And in the examples that we present, it is the structure that matters.

3.3 How to Use Background Knowledge

The next few pages will be devoted to an analysis of the interaction of background knowledge with a logical representation of a text. We will describe two modes of such an interaction; both seem to be present in our understanding of language. One exploits differences in plausibility of the meanings of words and phrases, in the absence of context (e.g., the difference between a central and a peripheral sense, or between a frequent and a rare meaning). The other one takes advantage of connections between those meanings. We do not claim that this is the only possible such analysis; rather,

we present a formal model which can perhaps be eventually disproved and replaced by a better one. As far as we know, this is the first such formal proposal.

3.3.1 Dominance. In Figure 1 the theories of “enter,” “ship,” etc. and the partial orders are represented graphically; more plausible theories are positioned higher. A path through this graph chooses an interpretation of the sentence S . For instance, the path $f_{int} = \{e1, sh1, p1, b1, d1\}$ and S say together that

A large boat (ship) that carries people or goods came into the harbor and carried a disease (illness).

Since it is the “highest” path, f_{int} is the most plausible (relative to R) interpretation of the words that appear in the sentence. Because it is also consistent, it will be chosen as a best interpretation of S , (cf. Zadrozny 1987a, 1987b). Another theory, consisting of $f' = \{e1, sh2, p1, b2, d1\}$ and S , saying that

A space vehicle came into the harbor and caused a disease/illness

is less plausible according to that ordering. As it turns out, f' is never constructed in the process of building an interpretation of a paragraph containing the sentence S , unless assuming f_{int} would lead to a contradiction, for instance within the higher level context of a science fiction story.

The collection of these most plausible consistent interpretations of a given theory T is denoted by $PT_{<}(T)$. Then f_{int} belongs to $PT_{<}(Th(\{S\}))$, but this is not true for f' .

Note: One should remember that, in general, because all our orderings are partial, there can be more than one most plausible interpretation of a sentence or a paragraph, and more that one “next best” interpretation. Moreover, to try to impose a total order on all the paths (i.e. the cartesian product defined in Section 3.3.2) would be a mistake; it would mean that ambiguities, represented in our formalism by existence of more than one (“best”) interpretation of a text, are outlawed.

3.3.2 Cartesian Products. Formalization of the interpretation produced in 3.3.1. is presented below. Any path through the graph of Figure 1 is an element of the cartesian product $\prod_{\Psi \in \text{subformulas}(S)} <_{\Psi}$ of the partial orderings.

Figure 2 explains the geometric intuitions we associate with the product and the ordering. The product itself is given by the following definition:

Definition

Let F be a collection of formulas $\psi_e, e \leq m$, for some natural number m ; and let, for each $e, <_e$ be a partial ordering on the collection theories

$$\{\psi_e \rightarrow T_0^e, \psi_e \rightarrow T_1^e, \dots, \psi_e \rightarrow T_{n_e}^e\}$$

of ψ_e . Define:

$$\Pi(F) = \prod_{e \leq m} <_e = \{f : (\forall e \leq m)(\exists l \leq n_e)[f(e) = \psi_e \rightarrow T_l^e]\}$$

We denote by $<$ the partial order induced on $\Pi(F)$ by the orderings $<_e$ and the canonical ordering of subformulas (a formula is “greater” than its subformulas). The geometrical meaning of this ordering can be expressed as “higher paths are more important provided they pass through the most specific nodes.”

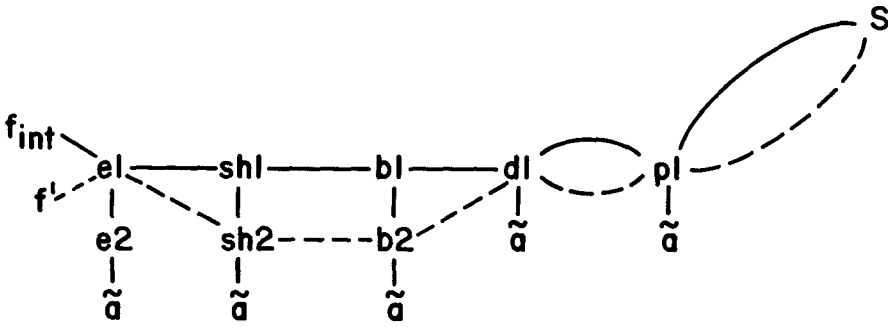


Figure 1
 The partial ordering of theories of the referential level **R** and the ordering of interpretations. Since (sh1) and (b1) dominate (respectively) (sh2) and (b2), the path f_{int} represents a more plausible interpretation than f' .

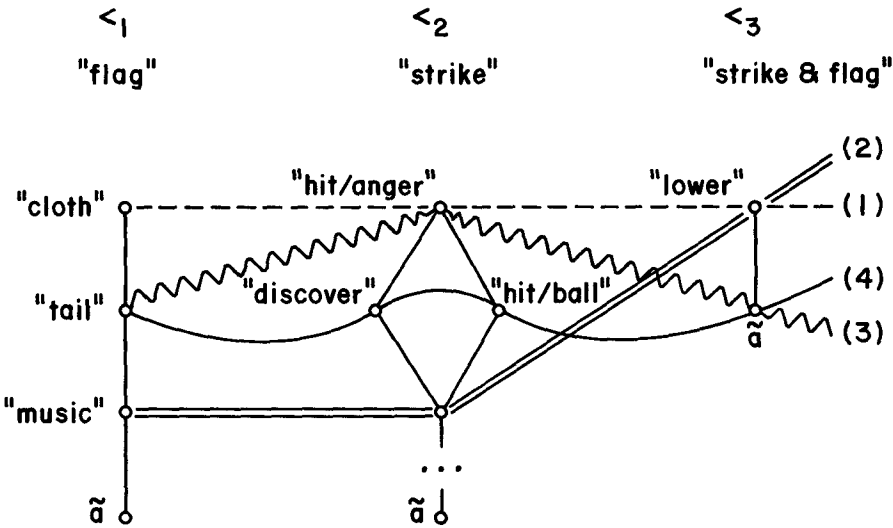


Figure 2
 The cartesian product $\prod_{i=1}^3 <_i$ can be depicted as a collection of all paths through the graphs representing the partial orderings; a path chooses only one element from each ordering—thus (1) and (2) are “legal” paths, while (4) is not. Also, more plausible theories appear higher: “cloth” > “music” > \tilde{a} . More specific paths are preferred: Assuming that all higher paths, like path (1), are excluded by inconsistency, path (2) is the most plausible interpretation of $\alpha \& \beta$, and it is preferred to (3). (More explanations in the text).

To make Figure 2 more intuitive we assigned some meanings to the partial orders. Thus, $<_1$ represents some possible meanings of “flag,” shown with the help of the “key words;” the meaning “piece of cloth” preferred to “deer’s tail.” The word “strike” has dozens of meanings, and we can imagine that the meaning of the transitive verb being represented by $<_2$, with “hit in anger” at the top, then “hit, e.g. a ball” and “discover” equally preferred, and then all other possible meanings. The trivial $<_3$ representing

“strike a flag” should remind us that we already know all that from Section 3.2. Notice that path (2) does not give us the correct interpretation of “strike the flag,” which is created from “cloth”- \tilde{a} -“lower.”

Each element of the cartesian product $\prod <_i$ represents a set of possible meanings. These meanings can be combined in various ways, the simplest of which consists of taking their union as we did in 3.3.1. But a paragraph isn't just a sum of its sentences, as a sentence isn't simply a concatenation of its phrases. The cohesion devices—such as “but,” “unless,” “since”—arrange sentences together, and they also have semantic functions. This is reflected, for instance, in the way various pieces of background knowledge are pasted together. Fortunately, at this point we can abstract from this by introducing an operator variable \oplus whose meaning will be, as a default, that of a set theoretic union, \cup ; but, as we describe it in Section 6.2, it can sometimes be changed to a more sophisticated join operator. There, when considering the semantics of “but,” we'll see that referential level theories can be combined in slightly more complicated ways. In other words, a partial theory corresponding to a paragraph cannot be just a sum of the theories of its sentences—the arrangement of those theories should obey the metalevel composition rules, which give the semantics of connectives. However, from a purely formal point of view, \oplus can be any function producing a theory from a collection of theories.

The cartesian product represents all possible amalgamations of these elementary theories. In other words, this product is the space of possible combinations of meanings, some of which will be inconsistent with the object level theory T . We can immediately exclude the inconsistent combinations, eliminating at least some nonsense:

$$\tilde{\Pi}(F) = \{f \in \Pi(F) : \oplus f \text{ is consistent with } T\}$$

It remains now to fill in the details of the construction of $PT_{<}$. We assume that a text P can be translated into a (ground) theory \vec{P} (a set of logical sentences); $T = Th(\vec{P})$ is the set of logical consequences of P . We denote by F the set $Form(Th(\vec{P}))$ —the set of all subformulas of $Th(\vec{P})$, about which we shall seek information at the referential level \mathbf{R} . If $F = \{\psi_1(\vec{c}_1), \dots, \psi_n(\vec{c}_n)\}$ (\vec{c}_i is a collection of constants that are arguments of ψ_i), is this theory, we have to describe a method of augmenting it with the background knowledge. We can assume without loss of generality that each $\psi_i(\vec{c}_i)$ in F has, in \mathbf{R} , a corresponding partial order $<_i$ of theories of $\psi_i(\vec{x}_i)$. We now substitute the constants \vec{c}_i for the variables \vec{x}_i inside the theories of $<_i$. With a slight abuse of notation, we will use the same symbol $<_i$ for the new ordering. The product spaces $\Pi(F)$ and $\tilde{\Pi}(F)$ can then be defined as before, with the new orderings in place of the ones with variables. Notice that if only some of the variables of $\psi_i(\vec{x}_i)$ were bound by \vec{c}_i , the same construction would work. We have arrived then at a general method of linking object level formulas with their theories from \mathbf{R} .

Now we can define $PT_{<}(T)$ of the theory T as the set of most likely consistent theories of T given by $(\Pi(F), <)$, where $F = Form(T)$:

$$PT_{<}(T) = \left\{ T \cup T' : T' = \oplus f \text{ and } f \text{ is a maximal element of } \left(\tilde{\Pi}(F), < \right) \right\}$$

Notice that $PT_{<}(T)$ can contain more than one theory, meaning that T is ambiguous. This is a consequence of the fact that the cartesian product is only partially ordered by $<$. The main reason for using ground instances $\psi_i(\vec{c}_i)$ in modifying the orderings is the need to deal with multiple occurrences of the same predicate, as in

John went to the bank by the bank.

The above construction is also very close in spirit to Poole's (1988) method for default reasoning, where object theories are augmented by ground instances of defaults.

3.3.3 Coherence Links. The reasoning that led to the intended interpretation f_{int} in our discussion of dominance was based on the partial ordering of the theories of \mathbf{R} . We want to exploit now another property of the theories of \mathbf{R} —their coherence. Finding an interpretation for a natural language text or sentence typically involves an appeal to coherence. Consider

S2: Entering the port, a ship brought a disaster.

Using the coherence link between (b2) and (dr1) (cf. Section 3.2)—the presence of *cause*(*,*) in the theories of "bring" and "disaster"—we can find a partial coherent interpretation $T \in PT_C(Th(\{S2\}))$ of *S2*. In this interpretation, theories explaining the meanings of terms are chosen on the basis of shared terms. This makes (b2) ("to bring" means "to cause") plausible and therefore it would be included in T . The formalization of all this is given below:

Definitions

- The set of all theories about the formulas of T is defined as:

$$\Sigma(T) = \{\psi \rightarrow T_i^\psi : \psi \in T \ \& \ \psi \rightarrow T_i^\psi \in <_\psi\}$$

Here, we ignore the ordering, because we are interested only in connections between concepts (represented by words).

- If $t, t' \in \Sigma(T)$, $t \neq t'$, share a predicate, we say that there is a **c-link** between t and t' . A **c-path** is defined as a chain of c-links; i.e. if $t = \psi \rightarrow T$ and $t' = \psi' \rightarrow T'$ belong to a c-path, then $\psi \neq \psi'$.
Under this condition, for any predicate, only one of its theories will belong to a c-path. A c-path therefore chooses only one meaning for each term.
- $C(T)$ will denote the set of all c-paths in $\Sigma(T)$ consistent with T , i.e. for each $p \in C(T)$, $\oplus p \cup T$ is consistent.
This construction is like the one we have encountered when defining $\tilde{\Pi}(T)$. The details should be filled out exactly as before; we leave this to the reader.
- We define $PT_C(T)$ of a theory T as the set of most coherent consistent theories of T given by $C(T)$:

$$PT_C(T) = \{T \cup T' : T' = \oplus_p \text{ and } p \text{ is a } \subset - \text{ maximal element of } C(T)\}$$

Going back to *S2*, $PT_C(Th(S2))$ contains also the interpretation based on the coherence link between "ship" and "bring," which involves "carry." Based on the just-described coherence relations, we conclude that sentence *S2* is ambiguous; it has two interpretations, based on the two senses of "bring." Resolution of ambiguities involves factors beyond the scope of this section—for instance, Gricean maxims and topic (Section 6), or various notions of context (cf. Small et al. 1988). We will continue the

topic of the interaction of object level theories with background knowledge by showing how the two methods of using background knowledge can be combined.

3.3.4 Partial Theories. A finer interpretation of an object level theory T —its *partial theory*—is obtained by the iteration:

$$PT(T) = PT_{<}(PT_C(Th(T)))$$

PT is well defined after we specify that PT of a set of theories is the set of the PT s (for both $<$ and C):

$$PT_{\alpha}(\{T_1, T_2, \dots\}) = PT_{\alpha}(T_1) \cup PT_{\alpha}(T_2) \cup \dots$$

Notice that coherence does not decide between (e1) and (e2) given the above \mathbf{R} , but the iteration produces two theories of S_2 , both of which assert that the meaning of “ship entered” is “ship came.”

A ship/boat came into the harbor/port and caused/brought a disaster.

A ship/boat came into the harbor/port and carried/brought a disaster.

$PT(\{S_1\})$ contains only one interpretation based on f_{int} :

A ship/boat came into the harbor/port and carried/brought a disease.

Partial theories will be the main syntactic constructs in the subsequent sections. In particular, the p-models will be defined as some special models or partial theories of paragraphs.

3.4 Summary and Discussion

We have shown that finding an interpretation of a sentence depends on two graph-theoretical properties—coherence and dominance. Coherence is a purely “associative” property; we are interested only in the existence of links between represented concepts/theories. Dominance uses the directionality of the partial orders.

A partial theory $PT(T)$ of an object theory T corresponding to a paragraph is obtained by joining most plausible theories or sentences, collocations, and words of the paragraph. However, this simple picture must be slightly retouched to account for semantic roles of inter- and intra-sentential connectives such as “but,” and to assure consistency of the partial theory. These modifications have complicated the definitions a little bit. The above definitions capture the fact that even if, in principle, any consistent combination of the mini-theories about predicates can be extended to an interpretation, we are really interested only in the most plausible ones. The theory $PT(T)$ is called “partial” because it does not contain all knowledge about predicates—less plausible properties are excluded from consideration, although they are accessible should an inconsistency appear. Moreover, the partiality is related to the unutilized possibility of iterating the operator PT (cf. Section 4).

How can we now summarize what we have learned about the three logical levels? To begin with, one should notice that they are syntactically distinct. If object level theories are expressed by collections of first order formulas, metalevel definitions—e.g., to express as a default that \oplus is a set theoretical union—require another language,

such as higher order logic or set theory, where one can define predicates dealing with models, consistency, and provability. Even if all background knowledge were described, as in our examples, by sets of first order theories, because of the preferences and inconsistencies of meanings, we could not treat R as a flat database of facts—such a model simply would not be realistic. Rather, R must be treated as a separate logical level for these syntactic reasons, and because of its function—being a pool of possibly conflicting semantic constraints.

The last point may be seen better if we look at some differences between our system and KRYPTON, which also distinguishes between an object theory and background knowledge (cf. Brachman et al. 1985). KRYPTON's A-box, encoding the object theory as a set of assertions, uses standard first order logic; the T-box contains information expressed in a frame-based language equivalent to a fragment of FOL. However, the distinction between the two parts is purely functional—that is, characterized in terms of the system's behavior. From the logical point of view, the knowledge base is the union of the two boxes, i.e. a theory, and the entailment is standard. In our system, we also distinguish between the "definitional" and factual information, but the "definitional" part contains collections of mutually excluding theories, not just of formulas describing a semantic network. Moreover, in addition to proposing this structure of R , we have described the two mechanisms for exploiting it, "coherence" and "dominance," which are not variants of the standard first order entailment, but abduction.

The idea of using preferences among theories is new, hence it was described in more detail. "Coherence," as outlined above, can be understood as a declarative (or static) version of marker passing (Hirst 1987; Charniak 1983), with one difference: the activation spreads to theories that share a predicate, not through the IS-A hierarchy, and is limited to elementary facts about predicates appearing in the text.

The metalevel rules we are going to discuss in Section 6, and that deal with the Gricean maxims and the meaning of "but," can be easily expressed in the languages of set theory or higher order logic, but not everything expressible in those languages makes sense in natural language. Hence, putting limitations on the expressive power of the language of the metalevel will remain as one of many open problems.

4. Coherence of Paragraphs

We are now in a position to use the notion of the referential level in a formal definition of coherence and topic. Having done that, we will turn our attention to the resolution of anaphora, linking it with the provability relation (abduction) \vdash_{R+M} and a metarule postulating that a most plausible model of a paragraph is one in which anaphors have references. Since the example paragraph we analyze has only one connective ("and"), we can postpone a discussion of connectives until Section 6.

Building an interpretation of a paragraph does not mean finding *all* of its possible meanings; the implausible ones should not be computed at all. This viewpoint has been reflected in the definition of a partial theory as a most plausible interpretation of a sequence of predicates. Now we want to restrict the notion of a partial theory by introducing the formal notions of *topic* and *coherence*. We can then later (Section 5.2) define *p-models*—a category of models corresponding to paragraphs—as models of *coherent* theories that satisfy all metalevel conditions.

The partial theories pick up from the referential level the most obvious or the most important information about a formula. This immediate information may be insufficient to decide the truth of certain predicates. It would seem therefore that the iteration of the *PT* operation to form a closure is needed (cf. Zadrozny 1987b).

However, there are at least three arguments against iterating *PT*. First of all, iteration would increase the complexity of building a model of a paragraph; infinite iteration would almost certainly make impossible such a construction in real time. Secondly, the cooperative principle of Grice (1975, 1978), under the assumption that referential levels of a writer and a reader are quite similar, implies that the writer should structure the text in a way that makes the construction of his intended model easy for the reader; and this seems to imply that he should appeal only to the most direct knowledge of the reader. Finally, it has been shown by Groesser (1981) that the ratio of derived to explicit information necessary for understanding a piece of text is about 8:1; furthermore, our reading of the analysis of five paragraphs by Crothers (1979) strongly suggests that only the most direct or obvious inferences are being made in the process of building a model or constructing a theory of a paragraph. Thus, for example, we can expect that in the worst case only one or two steps of such an iteration would be needed to find answers to *wh*-questions.

Let *P* be a paragraph, let $\vec{P} = (S_1, \dots, S_n)$ be its translation into a sequence of logical formulas. The set of all predicates appearing in *X* will be denoted by $Pred(X)$.

Definition

Let *T* be a partial theory of a paragraph *P*. A sequence of predicates appearing in \vec{P} , denoted by T_p , is called a **topic** of the paragraph *P*, if it is a longest sequence satisfying the conditions (1) and (2) below:

1. For all "sentences" S_i , (a) or (b) or (c) holds:

(a) *Direct reference to the topic:*

$$T_p \subset Pred(S_i)$$

(b) *Indirect reference to the topic:*

$$\text{If } \psi \in Pred(S_i) \ \& \ (\psi \rightarrow T_\psi) \in \mathbf{T}, \text{ then } T_p \subset Pred(T_\psi)$$

(c) *Direct reference to a previous sentence:*

$$\text{If } \psi \in Pred(S_i) \ \& \ (\psi \rightarrow T_\psi) \in \mathbf{T} \text{ then } Pred(S_{i-1}) \cap Pred(\psi \rightarrow T_\psi) \neq \varnothing$$

2. Either (i) or (ii) is satisfied:

(i) *Existence of a topic sentence:* $T_p \subset Pred(S_i)$, for some sentence S_i ;

(ii) *Existence of a topic sentence:* a theory of T_p belongs to \mathbf{R} , i.e. if θ is the conjunction of predicates of T_p then $\theta \rightarrow T_\theta \in \mathbf{R}$, for some T_θ .

The last two conditions say that either the discussed concept (topic) already exists in the background knowledge or it must be introduced in a sentence. For instance, we can see that the sentence *The effect of the Black Death was appalling* can be assumed to be a topic sentence.

The first three conditions make the requirements for a collection of sentences to have a topic. Either every sentence talks about the topic (as, for instance, the first two sentences of the paragraph about the Black Death), or a sentence refers to the topic

through background knowledge—the topic appears in a theory about an entity or a relation of the sentence (in the case of *Within twenty-four hours of infection . . .*, “infection” can be linked to “disease”—cf. Sections 2 and 4.2), or else a sentence elaborates a fragment of the previous sentence (the theme *The effect of. . .* being developed in *In less than. . .*).

The definition allows a paragraph to have more than one topic. For instance, a paragraph consisting of

John thinks Mary is pretty. John thinks Mary is intelligent. John wants to marry her.

can be either about $\{John(j), Mary(m), think(j, m, pretty(m))\}$, or about John, Mary, and marrying. (Notice that the condition 2 (i) forbids us merging the two topics into a larger one). Thus paragraphs can be ambiguous about what constitutes their topics. The point is that they should have one.

It is also clear that what constitutes a topic depends on the way the content of paragraph sentences is represented. In the last case, if “pretty” were translated into a predicate, and not into a modifier of m (i.e. an operator), “John thinking about Mary” could not be a topic, for it wouldn’t be the longest sequence of predicates satisfying the conditions (1) and (2).

We’d like to put forward a hypothesis that this relationship between topics and representations can actually be useful: Because the requirement that a well-formed paragraph should have a topic is a very natural one (and we can judge pretty well what can be a topic and what can’t), we can obtain a new method for judging semantic representations. Thus, if the naive first order representation containing $pretty(m)$ as one of the formulas gives a wrong answer as to what is the topic of the above, or another, paragraph, we can reject it in favor of a (higher order) representation in which adjectives and adverbs are operators, not predicates, and which provides us with an intuitively correct topic. Such a method can be used in addition to the standard criteria for judging representations, such as elegance and ability to express semantic generalizations.

Definition

A partial theory $T \in PT(\vec{P})$ of the paragraph P is *coherent* iff the paragraph P has a topic.

A random permutation of just any sentences about a disease wouldn’t be coherent. But it would be premature to jump to the conclusion that we need more than just existence of a topic as a condition for coherence. Although it may be the case that it will be necessary in the future to introduce notions like “temporal coherence,” “deictic coherence,” or “causal coherence,” there is no need to start multiplying beings now. We can surmise that the random permutations we talk about would produce an inconsistent theory; hence, the temporal, causal, and other aspects would be dealt with by consistency. But of course at this point it is just a hypothesis.

An important aspect of the definition is that coherence has been defined as a property of representation—in our case, it is a property of a formal theory. The existence of the topic, the direct or indirect allusion to it, and anaphora (which will be addressed below) take up the issue of formal criteria for a paragraph definition, which was raised

by Bond and Hayes (1983) (cf. also Section 2.1). The question of paragraph length can probably be attended to by limiting the size of p-models, perhaps after introducing some kind of metric on logical data structures.

Still, our definition of coherence may not be restrictive enough: two collections of sentences, one referring to "black" (about black pencils, black pullovers, and black poodles), the other one about "death" (war, cancer, etc.), connected by a sentence referring to both of these, could be interpreted as one paragraph about the new, broader topic "black + death." This problem may be similar to the situation in which current formal grammars allow nonsensical but parsable collections of words (e.g., "colorless green ideas..."), while before the advent of Chomskyan formalisms, a sentence was defined as the smallest meaningful collection of words; Fowler (1965, p. 546) gives 10 definitions of a sentence.

It then seems worth differentiating between the creation of a new concept like "black + death," with a meaning given by a paraphrase of the example collection of sentences, and the acceptance of the new concept—storing it in **R**. In our case the concept "black + death," which does not refer to any normal experiences, would be discarded as useless, although the collection of sentences would be recognized as a strange, even if coherent, paragraph.

We can also hope for some fine-tuning of the notion of topic, which would prevent many offensive examples. This approach is taken in computational syntactic grammars (e.g. Jensen 1986); the number of unlikely parses is severely reduced whenever possible, but no attempt is made to define *only* the so-called grammatical strings of a language.

Finally, as the paragraph is a natural domain in which word senses can be reliably assigned to words or sentences can be syntactically disambiguated, larger chunks of discourse may be needed for precise assignment of topics, which we view as another type of disambiguation. Notice also that for coherence, as defined above, it does not matter whether the topic is defined as a longest, a shortest, or—simply—a sequence of predicates satisfying the conditions (1) and (2); the existence of a sequence is equivalent with the existence of a shortest and a longest sequence. The reason for choosing a longest sequence as the topic is our belief that the topic should rather contain more information about a paragraph than less.

4.1 Comparison with Other Approaches

At this point it may be proper to comment on the relationship between our theory of coherence and theories advocated by others. We are going to make such a comparison with the theories proposed by J. Hobbs (1979, 1982) that represent a more computationally oriented approach to coherence, and those of T.A. van Dijk and W. Kintch (1983), who are more interested in addressing psychological and cognitive aspects of discourse coherence. The quoted works seem to be good representatives for each of the directions; they also point to related literature.

The approach we advocate is compatible with the work of these researchers, we believe. There are, however, some interesting differences: first of all, we emphasize the role of paragraphs; second, we talk about formal principles regulating the organization and use of knowledge in language understanding; and third, we realize that natural language text (such as an on-line dictionary) can, in many cases, provide the type of commonsense background information that Hobbs (for example) advocated but didn't know how to access. (There are also some other, minor, differences. For instance, our three-level semantics does not appeal to possible worlds, as van Dijk and Kintch do; neither is it objectivist, as Hobbs' semantics seems to be.)

We shall discuss only the first two points, since the third one has already been explained.

The chief difference between our approach and the other two lies in identifying the paragraph as a domain of coherence. Hobbs, van Dijk, and Kintch distinguish between "local" coherence—a property of subsequent sentences—and "global" coherence—a property of discourse as a whole. Hobbs explains coherence in terms of an inventory of "local," possibly computable, coherence relations, like "elaboration," "occasion," etc. (Mann and Thompson 1983 give an even more detailed list of coherence relations than Hobbs.) Van Dijk and Kintch do this too, but they also describe "macrostructures" representing the global content of discourse, and they emphasize psychological and cognitive strategies used by people in establishing discourse coherence. Since we have linked coherence to models of paragraphs, we can talk simply about "coherence"—without adjectives—as a property of these models. To us the first "local" domain seems to be too small, and the second "global" one too large, for constructing meaningful computational models. To be sure, we believe relations between pairs of sentences are worth investigating, especially in dialogs. However, in written discourse, the smallest domain of coherence is a paragraph, very much as the sentence is the basic domain of grammaticality (although one can also judge the correctness of phrases).

To see the advantage of assuming that coherence is a property of a fragment of a text/discourse, and not a relation between subsequent sentences, let us consider for instance the text

John took a train from Paris to Istanbul. He likes spinach.

According to Hobbs (1979, p. 67), these two sentences are incoherent. However, the same fragment, augmented with the third sentence *Mary told him yesterday that the French spinach crop failed and Turkey is the only country...* (ibid.) suddenly (for Hobbs) becomes coherent. It seems that any analysis of coherence in terms of the relation between subsequent sentences cannot explain this sudden change; after all, the first two sentences didn't change when the third one was added. On the other hand, this change is easily explained when we treat the first two sentences as a paragraph: if the third sentence is not a part of the background knowledge, the paragraph is incoherent. And the paragraph obtained by adding the third sentence is coherent. Moreover, coherence here is clearly the result of the existence of the topic "John likes spinach."

We derive coherence from formal principles regulating the organization and use of knowledge in language understanding. Although, like the authors discussed above, we stress the importance of inferencing and background knowledge in determining coherence, we also address the problem of knowledge organization; for us the central problem is how a model emerges from such an organization. Hobbs sets forth hypotheses about the interaction of background knowledge with sentences that are examined at a given moment; van Dijk and Kintch provide a wealth of psychological information on that topic. But their analyses of how such knowledge could be used are quasi-formal. Our point of departure is different: we assume a certain simple structure of the referential level (partial orders) and a natural way of using the knowledge contained there ("coherence links" + "most plausible = first"). Then we examine what corresponds to "topic" and "coherence"—they become **mathematical** concepts. In this sense our work refines these concepts, changes the way of looking at them by linking them to the notion of paragraph, and puts the findings of the other researchers into a new context.

5. Models of Paragraphs

We argue below that paragraphs can be mapped into models with small, finite universes. We could have chosen another, more abstract semantics, with infinite models, but in this and all cases below we have in mind computational reasons for this enterprise. Thus, as in the case of Kamp's (1981) DRS, we shall construct a kind of Herbrand model of texts, with common and proper names translated into unary predicates, intransitive verbs into unary predicates, and transitive verbs into binary predicates. In building the logical model M of a collection of formulas \vec{S} corresponding to the sentences of a paragraph, we assume that the universe of M contains constants introduced by elements of S , usually by ones corresponding to NPs, and possibly by some formulas picked by the construction from the referential level. However, we are interested not in the relationship between truth conditions and representations of a sentence, but in a formalization of the way knowledge is used to produce a representation of a section of text. Therefore we need not only a logical description of the truth conditions of sentences, as presented by Kamp, but also a formal analysis of how background knowledge and metalevel operations are used in the construction of models. This extension is important and nontrivial; we doubt that one can deal effectively with coherence, anaphora, presuppositions or the semantics of connectives without it. We have begun presenting such an analysis in Section 3, and we continue now.

5.1 The Example Revisited: Preparation for Building a Model

We return now to the example paragraph, to illustrate how the interaction between an object theory and a referential level produces a coherent interpretation of the text (i.e., a p -model) and resolves the anaphoric references. The method will be similar to, but more formal than, what was presented in Section 2. In order not to bore the reader with the same details all over again, we will use a shorter version of the same text.

Example 2

- P₁: In 1347 a ship entered the port of Messina bringing with it the disease that came to be known as the Black Death.
- P₂: It struck rapidly.
- P₃: Within twenty-four hours of infection came an agonizing death.

5.1.1 Translation to Logic. The text concerns events happening in time. Naturally, we will use a logical notation in which formulas may have temporal and event components. We assume that any formal interpretation of time will agree with the intuitive one. So it is not necessary now to present a formal semantics here. The reader may consult recent papers on this subject (e.g. Moens and Steedman 1987; Webber 1987) to see what a formal interpretation of events in time might look like. Since sentences can refer to events described by other sentences, we may need also a quotation operator; Perlis (1985) describes how first order logic can be augmented with such an operator. Extending and revising Jackendoff's (1983) formalism seems to us a correct method to achieve the correspondence between syntax and semantics expressed in the *grammatical constraint* ("that one should prefer a semantic theory that explains otherwise arbitrary generalizations about the syntax and the lexicon"—*ibid.*).

However, as noted before, we will use a simplified version of such a logical notation; we will have only **time**, **event**, **result**, and **property** as primitives. After these remarks we can begin constructing the model of the example paragraph. We assume

that constants are introduced by NPs. We have then

(i) **Constants** $s, m, d, i, b, 1347$ satisfying: $ship(s), Messina(m), disease(d), infection(i), death(b), year(1347)$.

(ii) **Formulae**

S_1 : [**time** : $year(1347)$; **event** : $enter(s, m) \ \& \ ship(s) \ \& \ port(m) \ \& \ bring(x_0, d) \ \& \ disease(d) \ \& \ name(d, \mathbf{BlackDeath}) \ \& \ (x_0 = s \vee x_0 = d \vee x_0 = m)$]

S_2 : **time** : $past$; **event** : $rapidly : strike(y_0) \ \& \ (y_0 = s \vee y_0 = m \vee y_0 = d)$]

S_3 : $\exists t, t' \{$ **time** : $t; infection(i)$ $\} \ \& \ [$ **time**: $t' \in (t, t + 24h)$ $;$
event : $come(b) \ \& \ death(b) \ \& \ agonizing(b)$ $\}]$

The notation **time** : $\alpha(t)$; **event** : β should be understood as meaning that the event described by the formula β took place in (or during) the time period described by the formula $\alpha(t)$, t ranges over instants of time (not intervals).

Note. We assume that “strike” is used intransitively. But our construction of the p-models of the paragraph would look exactly the same for the transitive meaning, except that we would be expected to infer that people were harmed by the illness.

5.1.2 Referential Level. We use only standard dictionaries as a source of background information. The content of this information is of secondary importance—we want to stress the formal, logical side of the interaction between the referential level and the object theory. Therefore we represent both in our simplified logical notation, and not in English. All formulas at the referential level below have been obtained by a direct translation of appropriate entries in Webster’s and Longman. The translation in this case was manual, but could be automated.

- **Referential level** (a fragment):

$$\psi \rightarrow T_\psi$$

$ship(x) \rightarrow \{large : boat(x); \exists y carry(x, y) \ \& \ (people(y) \vee goods(y)) \ \& \ agent(x); \dots\}$ (sh1)

/* ship—a large boat for carrying people or goods on the sea */

$bring(x, y) \rightarrow \{carry(x, y); \dots\}$ (b1)

/* bring—to carry */

$strike(x, y) \rightarrow \{hit(x, y); agent(x) \ \& \ patient(y); \dots\}$ (s1a)

$strike(x) \rightarrow \{hit(x); agent(x); \dots\}$ (s1b)

/* strike—to hit */

$strike(x) \rightarrow \{illness(x) \ \& \ \exists y suddenly : harm(x, y); \dots\}$ (s2.ex)

/* strike—to harm suddenly; “they were struck by illness */

- $disease(y) \rightarrow \{illness(y) \ \& \ \exists z (infection(z) \ \& \ causes(z, y)); \dots\}$ (d1)
 /* disease—illness caused by an infection */
- $death(x) \rightarrow \{\exists t, y[x = ' [time : t; event : die(y) \ \& \ (creature(y) \vee plant(y))]']\}$ (de.1)
 /* death—an event in which a creature or a plant dies */
- $come(x) \rightarrow \{\exists t[time : t; event : arrive(x)]\}$ (ct.1)
 /* to come—to arrive (...) in the course of time */
- $infection(x) \rightarrow \{\exists e, y, z[e = event : infect(y, z) \ \& \ person(y) \ \& \ disease(z)]\}$
 & $x = result(e)$ (i.1)
 /* infection—the result of being infected by a disease */
- $agonizing(x) \rightarrow \{\exists y causes(x, y) \ \& \ pain(y)\}$ (a.1)
 /* agonizing—causing great pain */
- $enter(x, y) \rightarrow \{come_in(x, y); place(y)\}$ (e.1)
 /* enter—to come into a place */
- ...

We have shown, in Section 3, the role of preferences in building the model of a paragraph. Therefore, to make our exposition clearer, we assume that all the above theories are equally preferred. Still, some interesting things will happen before we arrive at our intended model.

5.1.3 Provability and Anaphora Resolution. To formalize a part of the process for finding antecedents of anaphors, we have to introduce a new logical notion—the relation of *weak R + M-abduction*. This relation would hold, for instance, between the object theory of our example paragraph and a formula expressing the equality of two constants, i and i' , denoting (respectively) the “infection” in the sentence *Within twenty-four hours of infection...*, and the “infection” of the theory (d1)—*a disease is an illness caused by an infection*. This equality $i = i'$ cannot be proven, but it may be reasonably assumed—we know that in this case the infection i' caused the illness, which, in turn, caused the death.

The necessity of this kind of merging of arguments has been recognized before: Charniak and McDermott (1985) call it *abductive unification/matching*, Hobbs (1978, 1979) refers to such operations using the terms *knitting* or *petty conversational implicature*. Neither Hobbs nor Charniak and McDermott tried then to make this notion precise, but the paper by Hobbs et al. (1988) moves in that direction. The purpose of this subsection is to formalize and explain how assumptions like that one above can be made.

Definition

A formula ϕ is *weakly provable* from an object theory T , expressed as $T \vdash_R \phi$, iff there exists a partial theory $T \in PT(T)$ such that $T \vdash \phi$, i.e. T proves ϕ in logic. (We call \vdash_R “weak” because it is enough to find **one** partial theory proving a given formula.)

As an example, in the case of the three-sentence paragraph, we have a partial theory T_1 based on (s1b) saying that “it **hits** rapidly,” and T_2 saying that “an illness (‘it’) **harms** rapidly” (s2.ex). Thus both statements are weakly provable.

Since we view the metalevel constraints \mathbf{M} rather as rules for choosing models than as special inference rules, the definition of the $\mathbf{R}+\mathbf{M}$ -abduction is model-theoretic, not proof-theoretic:

Definition

A *preferred model* of a theory T is an element of $Mods(T)$ that satisfies metalevel constraints contained in \mathbf{M} . The set of all preferred models of T is denoted by $PM(T)$.

A formula ϕ of $L(=)$, the language with equality, is weakly $R + M$ -abductible from an object theory T , denoted by $T \vdash_{R+M} \phi$, iff there exists a partial theory $T \in PT(T)$ and a preferred model $M \in PM(T)$ such that $M \models \phi$, i.e. ϕ is true in at least one preferred model of the partial theory T .

Note: The notions of strong provability and strong $R + M$ -abduction can be introduced by replacing “there exists” by “all” in the above definitions (cf. Zadrozny 1987b). We will have, however, no need for “strong” notions in this paper. Also, in a practical system, “satisfies” should be probably replaced by “violates fewest.”

Obviously, it is better to have references of pronouns resolved than not. After all, we assume that texts make sense, and that authors know these references. That applies to references of noun phrases too. On the other hand, there must be some restrictions on possible references; we would rather assume that “spinach” \neq “train” (i.e. $\forall x, y)(spinach(x) \ \& \ train(y) \rightarrow x \neq y)$), or “ship” \neq “disease.” Two elementary conditions limiting the number of equalities are: an equality $N1 = N2$ may be assumed only if either $N1$ and $N2$ are listed as synonyms (or paraphrases) or their equality is explicitly asserted by the partial theory T . Of course there are other conditions, like “typically, the determiner ‘a’ introduces a new entity, while ‘the’ refers to an already introduced constant.” (But notice that in our example paragraph “infection” appears without an article.) All these, and other, guidelines can be articulated in the form of metarules.

We define another partial order, this time on models $Mods(T)$ of a partial theory T of a paragraph: $M_1 \geq M_2$, if M_1 , satisfies more $R + M$ -abductible equalities than M_2 . The principle articulating preference for having the references resolved can now be expressed as

Metarule 1

Assume that $T \in PT(\vec{P})$ is a partial theory of a paragraph P . Every preferred model $M \in PM(T)$ is a maximal element of the ordering \geq of $Mods(T)$.

To explain the meaning of the metarule, let us analyze the paragraph (P_1, P_2, P_3) and the background knowledge needed for some kind of rudimentary understanding of that text. The rule (i.1) (*infection is a result of being infected by a disease...*), dealing with the infection i , introduces a disease d_1 ; we also know about the existence of the disease d in 1347. Now, notice that there may be many models satisfying the object theory of the paragraph P augmented by the background knowledge. But we can find two among them: in one, call it M_1 , d and d_1 are identical; in the other one, M_2 , they are distinct. The rule says that only the first one has a chance to be a preferred model of the paragraph; it has more noun phrase references resolved than the other model, or—formally—it satisfies more $R + M$ -abductible equalities, and therefore $M_1 \geq M_2$.

This reasoning, as the reader surely has noticed, resembles the example about infections from the beginning of this section. The difference between the cases lies in the equality $d = d_1$ being the result of a formal choice of a model, while $i = i'$ wasn't proved, just “reasonably” assumed.

In interpreting texts, knowledge of typical subjects and typical objects of verbs helps in anaphora resolution (cf. Braden-Harder & Zadrozny 1990). Thus if we know that *A farmer grows vegetables*, either having obtained this information directly from a text, or from *R*, we can reasonably assume that *He also grows some cotton* refers to the farmer, and not to a policeman mentioned in the same paragraph. Of course, this should be only a defeasible assumption, if nothing indicates otherwise. We now want to express this strategy as a metarule:

Metarule 2

Let us assume that it is known that $P(a, b) \ \& \ Q(a) \ \& \ R(b)$, and it is not known that $P(a', X)$, for any X . Then models in which $P(a, c) \ \& \ R'(c)$ holds are preferred to models in which $P(a', c) \ \& \ R'(c)$ is true.

One can think of this rule as a model-theoretic version of Ockham's razor or abduction; it says "minimize the number of things that have the property $P(*, *)$," and it allows us to draw certain conclusions on the basis of partial information. We shall see it in action in Section 5.2.

We have no doubts that various other metarules will be necessary; clearly, our two metarules cannot constitute the whole theory of anaphora resolution. They are intended as an illustration of the power of abduction, which in this framework helps determine the universe of the model (that is the set of entities that appear in it). Other factors, such as the role of focus (Grosz 1977, 1978; Sidner 1983) or quantifier scoping (Webber 1983) must play a role, too. Determining the relative importance of those factors, the above metarules, and syntactic clues, appears to be an interesting topic in itself.

Note: In our translation from English to logic we are assuming that "it" is anaphoric (with the pronoun following the element that it refers to), not cataphoric (the other way around). This means that the "it" that brought the disease in P_1 will not be considered to refer to the infection "i" or the death "d" in P_3 . This strategy is certainly the right one to start out with, since anaphora is always the more typical direction of reference in English prose (Halliday and Hasan 1976, p. 329).

Since techniques developed elsewhere may prove useful, at least for comparison, it is worth mentioning at this point that the proposed metarules are distant cousins of "unique-name assumption" (Genesereth and Nilsson 1987), "domain closure assumption" (ibid.), "domain circumscription" (cf. Etherington and Mercer 1987), and their kin. Similarly, the notion of $R + M$ -abduction is spiritually related to the "abductive inference" of Reggia (1985), the "diagnosis from first principles" of Reiter (1987), "explainability" of Poole (1988), and the subset principle of Berwick (1986). But, obviously, trying to establish precise connections for the metarules or the provability and the $R + M$ -abduction would go much beyond the scope of an argument for the correspondence of paragraphs and models. These connections are being examined elsewhere (Zadrozny forthcoming).

5.2 p-Models

The construction of a model of a paragraph, a *p-model*, must be based on the information contained in the paragraph itself (the object theory) and in the referential level, while the metalevel restricts ways that the model can be constructed, or, in other words, provides criteria for choosing a most plausible model(s), if a partial theory is ambiguous. This role of the metarules will be clearly visible in finding references of pronouns in a simple case requiring only a rule postulating that these references be

searched for, and in a more complex case (in Section 5) when they can be found only by an interplay of background knowledge and (a formalization of) Gricean maxims.

Definition

M is a *p-model* of a paragraph P iff there exists a coherent partial theory $T \in PT(\vec{P})$ such that $M \in PM(T)$.

Having defined the notion of a p-model, we can mimic now, in logic, the reasoning presented in Section 2.2. Using background information and the translation of sentences, we build a p-model of the paragraph. This involves determining the references of the pronoun “it,” and deciding whether “struck” in the sentence *It struck rapidly* means “hit” (s1b) or “harmed” (s2.ex). We have then two meanings of “strike” and a number of possibilities for the pronouns.

We begin by constructing the two classes of preferred models given by (s1b) and (s2.ex), respectively. It is easily seen that, in the models of the first class, based on $\{S_2, s1b\}$, (that is $\{rapidly : strike(y_0)$ and $strike(x) \rightarrow hit(x) \dots\}$), together with all other available information, do not let us $R+M$ -abduct anything about y_0 , i.e., the referent for the subject pronoun “it” in P_2 (*it struck rapidly*). On the other hand, from $\{S_2, s2.ex, d1\}$ we $R+M$ -abduct that $y_0 = d$, i.e. *the disease struck rapidly*. That is the case because *s2.ex* implies that the agent that “struck rapidly” is actually an illness. From *rapidly : strike*(y_0), *strike*(x) \rightarrow *illness*(x) & \dots , *disease*(y) \rightarrow *illness*(y) & \dots , and *disease*(d) we can infer *illness*(y_0) and *illness*(d); by the Metarule (1) we conclude that $y_0 = d$. In other words, the referent for the subject “it” is “disease.” Thus the Metarule (1) immediately eliminates all the models from the first class given by (s1b), in which “struck” means “hit.”

Notice that we cannot prove in classical logic that the ship has brought the disease. But we are allowed to assume it by the above formal rule as the most plausible state of affairs, or—in other words—we prove it in our three-level logic.

We are left then with models of the three sentences (S_1, S_2, S_3) that contain $\{S_2, s2.ex, d1\}$; they all satisfy $y_0 = d$. We now use $\{S_1, sh1, b1\}$ (*enter*(s, m) & *ship*(s) & *bring*(x_0, d) & \dots ; *ship*(s) \rightarrow ($\exists y$)*carry*(s, y) & \dots ; $\forall z$ [*bring*(x_0, z) \rightarrow *carry*(x_0, z)]). From these facts we can conclude by Metarule (1) that $x_0 = s$: a “ship” is an agent that carries goods; to “bring” means to “carry”; and the disease has been brought by something—we obtain *carry*(x_0, d) and *carry*(s, y); and then by Metarule (2), *carry*(s, d). That is, the referent for the pronoun “it” in P_1 (\dots *bringing with it the disease* \dots) should be “ship.” Observe that we do not assert about the disease that it is a kind of goods or people; the line of reasoning goes as follows: since ships are known to carry people or goods, and ports are not known to carry anything, we may assume that the ship carried the disease along with its standard cargo.

Having resolved all pronoun references, with no ambiguity left, we conclude that the class $PM(P)$ consists of only one model, based on the the partial theory

$$\{S_1, S_2, S_3, sh1, b1, e.1, s2.ex, d1, de1, i1, a1, ct1\}.$$

The model describes a situation in which the ship came into the port/harbor; the ship brought the disease; the disease was caused by an infection; the disease harmed rapidly, causing a painful death; and so on.

The topic T_p of (P_1, P_2, P_3) is the *disease*(x). The first sentence talks about it; the second one refers to it using the pronoun “it,” and the third one extends our knowledge about the topic, since “disease” is linked to “infection” through $d1$. Furthermore,

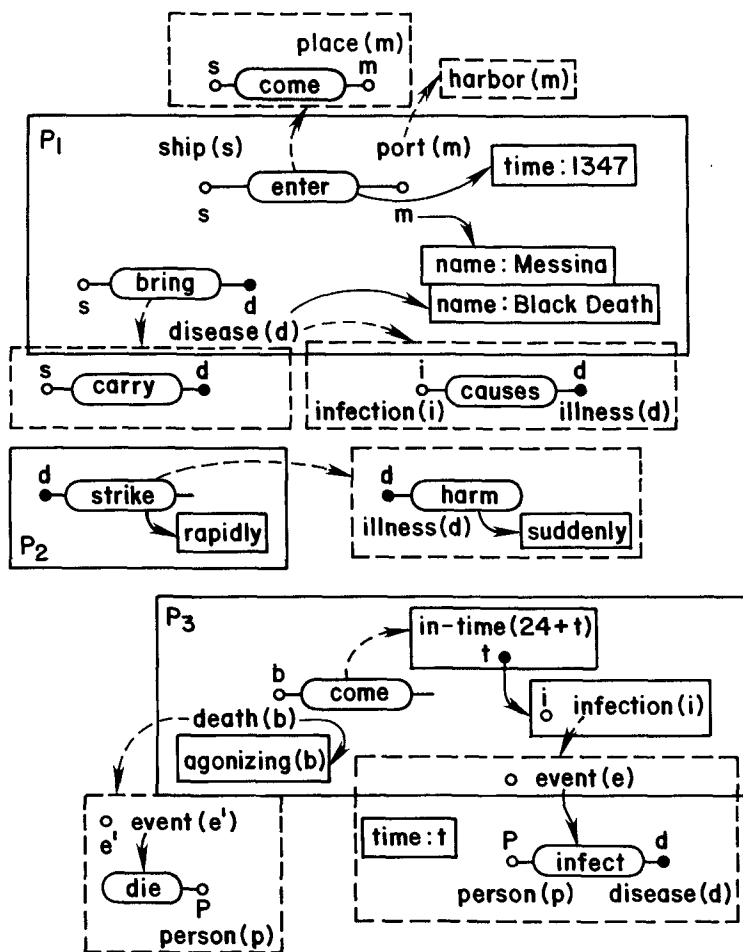


Figure 3
The p-model for the example paragraph

“disease” is the only noun phrase mentioned or referred to in all three sentences; the sentence S_1 is the topic sentence. The p-model of the paragraph is represented by Figure 3.

But notice that our definition of topic licences also other analyses, for example, one in which all the predicates of the first sentence constitute the topic of the paragraph, S_2 elaborates S_1 (in the sense of condition 1 (c) of the definition of topic), and S_3 elaborates S_2 . Based on the larger context, we prefer the first analysis; however, a computational criterion for such a preference remains as an open problem.

6. On the Role of the Metalevel

We have already seen examples of the application of metalevel rules. In the analysis of the paragraph, we applied one such rule expressing our commonsense knowledge about the usage of pronouns. In this section we discuss two other sources of metalevel axioms: Gricean cooperative principles, which reduce the number of possible

interpretations of a text or an utterance; and connectives and modalities—such as “but,” “unless,” or “maybe”—which refer to the process of constructing the models or partial theories, and to some operations on them (see Figure 4).

We can see then two applications of metarules: in constructing models of a text from representations of sentences, and in reducing, or constraining, the ambiguity of the obtained structure. We begin by showing how to formalize the latter. In the next subsection (6.1), assuming the Gricean maxims to be constraints on language communication, either spoken or written, we use their formal versions in building partial theories. A specific instance of the rule of “quantity” turns out to be applicable to anaphora resolution. That example will end our discussion of anaphora in this article.

The last topic we intend to tackle is the semantic role of conjunctions. In subsection 6.2 we present a metalevel axiom dealing with the semantic role of the adversative conjunction “but;” then we talk about some of its consequences for constructing models of text. This will complete our investigation of the most important issues concerning paragraph structure: coherence (how one can determine that a paragraph expresses a “thought”), anaphora (how one can compute “links” between entities that a paragraph talks about), and cohesion (what makes a paragraph more than just a sum of sentences). Of course, we will not have final answers to any of these problems, but we do believe that the/a direction of search for computational models of text will be visible at that point.

We assume a flat structure of the metalevel, envisioning it as a collection of (closed) formulas written in the language of set theory or higher order logic. In either of the two theories it is possible to define the notions of a model, satisfiability, provability, etc. for any first order language (cf. e.g. Shoenfield 1967); therefore the metalevel formulas can say how partial theories should be constructed (specifying for instance the meaning of \oplus) and what kinds of models are admissible. The metarules thus form a logical theory in a special language, such as the language of ZF-set theory. However, for the sake of readability, we express all of them in English.

6.1 A Formalization of Gricean Maxims

A Gricean Cooperative Principle applies to text, too. For instance, in normal writing people do not express common knowledge about typical functions of objects. In fact, as the reader may check for himself, there is nothing in Gricean maxims that does not apply to written language. That the maxims play a semantic role is hardly surprising. But that they can be axiomatized and used in building formal models of texts is new. We present in the next couple of paragraphs our formalization of the first maxim, and sketch axiomatizations of the others. Then we will apply the formal rule in an example.

Gricean maxims, after formalization, belong to the metalevel. This can be seen from our formalization of the rule “don’t say too much.” To this end we define *redundancy* of a partial theory T of a paragraph as the situation in which some sentences can be logically derived from other sentences and from the theory T in a direct manner:

$$(\exists S \in \vec{P})(\exists \sigma \in \mathbf{R})[\sigma \in T \ \& \ \sigma = \psi \rightarrow \phi \ \& \ \vec{P} \vdash \psi \ \& \ \{\phi\} \cup (\vec{P} - \{S\}) \vdash S]$$

The meaning of this formula can be explained as follows: a paragraph P has been translated into its formal version \vec{P} and is to be examined for redundancy. Its partial theory $PT(\vec{P})$ has also been computed. The test will turn positive if, for some sentence S , we can find a rule/theorem $\sigma = \psi \rightarrow \phi$ in $PT(\vec{P})$ such that the sentence S is implied (in a classical logic) by the other sentences and σ . For example, if the paragraph about Black Death were to contain also the sentence *The ship carried people or goods, or*

- Quantity. Say neither too much nor too little.
- Quality. Try to make your contribution one that is true.
- Relation. Be relevant.
- Manner. Avoid obscurity and ambiguity; be brief and orderly.

Figure 4

The Gricean maxims

both, which (in its logical form) belongs to **R**, it would be redundant: $\sigma = (sh1)$, there. Similarly, the definition takes care of the redundancy resulting from a simple repetition.

Metarule G1a

(*nonredundancy*) If $T_1, T_2 \in PT(\vec{P})$ and T_1 is less redundant than T_2 , then the theory T_1 is preferred to T_2 . (Where “less redundant” means that the number of redundant sentences in T_1 is smaller than in T_2)

The relevant half of the Maxim of Quantity has been expressed by **G1a**. How would we express the other maxims? The “too little” part of the first maxim might be represented as a preference for unambiguous partial theories. The second maxim has been assumed all the time—when constructing partial theories or models, the sentences of a paragraph are assumed to be true. The Maxim of Manner seems to us to be more relevant for critiquing the style of a written passage or for natural language generation; in the case of text generation, it can be construed as a requirement that the produced text be coherent and cohesive.

We do not claim that **G1a** is the best or unique way of expressing the rule “assume that the writer did not say too much.” Rather, we stress the possibility that one can axiomatize and productively use such a rule. We shall see this in the next example: two sentences, regarded as a fragment of paragraph, are a variation on a theme by Hobbs (1979).

Example 3

The captain is worried because the third officer can open his safe. He knows the combination.

The above metarule postulating “nonredundancy” implies that “he” = “the third officer,” “his” = “the captain’s” are the referents of the pronouns. This is because the formula

$$safe(x) \rightarrow (owns(y, x) \ \& \ cmbntn(z, x) \rightarrow knows(y, z) \ \& \ can_open(y, x)) \in T_{safe},$$

belongs to **R**, since it is common knowledge about safes that they have owners, and also combinations that are known to the owners. Therefore “his” = “the third officer’s” would produce a redundant formula, corresponding to the sentence *The third officer can open the third officer’s safe*. By the same token, *The captain knows the combination* would be redundant too.

We now explain the details of this reasoning. One first proves that “his” = “the captain’s.” Indeed, if “his” = “the third officer’s,” then our example sentence would mean

? *The captain is worried because the third officer can open the third officer’s safe;*

in logic:

$captain(x) \ \& \ worry(x,s) \ \& \ sentence(s)$
 $\& \ s = \text{‘the third officer can open the third officer’s safe.’}$

We assume also, based on common knowledge about worrying, that $worry(x', s') \rightarrow S$. That is, one worries about things that might possibly be or become true (S denotes the logical formula corresponding to the sentence s , cf. Section 3); but one doesn’t worry about things that are accepted as (almost) always true (such as the law of gravity), so that $worry(x', s') \rightarrow \neg(S \in T_f)$, where f ranges over subformulas of S .

In our case, S immediately follows from T_{safe} and X , where $X = safe(sf) \ \& \ third.officer(o) \ \& \ owns(o, sf)$ —the fact that “the third officer can open the third officer’s safe” is a consequence of general knowledge about the ownership of safes. And therefore the interpretation with “his” = “the captain’s” is preferred as less redundant by the rule **G1a**. This theory contains representations of the two sentences, the theory of safes, a theory of worrying, and the equality “his” = “captain’s.”

It remains to prove that “he” = “the third officer.” Otherwise we have

P1. *The captain is worried because the third officer can open the captain’s safe.*

P2. ? *The captain knows the combination.*

Clearly, the last sentence is true but redundant—the theory of “safe” and P1 entail P2:

$$\{P1\} \cup T_{safe} \vdash P2$$

We are left with the combination

Q1. *The captain is worried because the third officer can open the captain’s safe.*

Q2. ? *The third officer knows the combination.*

In this case, Q2 does not follow from $\{Q1\} \cup T_{safe}$ and therefore Q1, Q2 is preferred to P1, P2 (by **G1a**). We obtain then

The captain is worried because the third officer can open the captain’s safe.

The third officer knows the combination

as the most plausible interpretation of our example sentences.

Note: The reader must have noticed that we did not bother to distinguish the sentences P1, P2, Q1 and Q2 from their logical forms. Representing “because” and “know” adequately should be considered a separate topic; representing the rest (in the first order convention of this paper) is trivial.

6.1.1 Was the Use of a Gricean Maxim Necessary? Can one deal effectively with the problem of reference without axiomatized Gricean maxims, for instance by using only

“petty conversational implicature” (Hobbs 1979), or the metarules of Section 5.2? It seems to us that the answer is no.

As a case in point, consider the process of finding the antecedent of the anaphor “he” in the sentences

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

Hobbs (1979, 1982) proves “he” = “John” by assuming the relation of “elaboration” between the sentences. (Elaboration is a relation between two segments of a text. It intuitively means “expressing the same thought from a different perspective,” but has been defined formally as the existence of a proposition implied by both segments—here the proposition is “John can open the safe.”) However, if we change the pair to the triple

Bill has a safe under the painting of his yacht. John can open Bill's safe. He knows the combination

the relation of elaboration holds between the segment consisting of the first two sentences of the triple and each of the two possible readings: *John knows the combination* and *Bill knows the combination*. In this case, elaboration cannot choose the correct referent, but the rule G1a can and does. Clearly, an elaboration should not degenerate into redundancy; the Gricean maxims are to keep it fresh.

As we have observed, correct interpretations cannot be chosen by an interaction of an object level theory and a referential level alone, because coherence, plausibility and consistency are too weak to weed out wrong partial theories. Metarules are necessary. True, the captain knew the combination, but it was consistent that “his” might have referred to “the third officer’s.”

6.2 Semantics of the Conjunction “But”

Any analysis of natural language text, to be useful for a computational system, will have to deal with coherence, anaphora, and connectives. We have examined so far the first two concepts; we shall present now our view of connectives to complete the argument about paragraphs being counterparts of models. We present a metalevel rule that governs the behavior of the conjunction “but;” we formalize the manner in which “but” carries out the contradiction. Then we derive from it two rules that prevent infelicitous uses of “but.”

Connectives are function words—like conjunctions and some adverbs—that are responsible simultaneously for maintaining cohesiveness within the text and for signaling the nature of the relationships that hold between and among various text units. “And,” “or,” and “but” are the three main coordinating connectives in English. However, “but” does not behave quite like the other two—semantically, “but” signals a contradiction, and in this role it seems to have three subfunctions:

1. Opposition (called “adversative” or “contrary-to-expectation” by Halliday and Hasan 1976; cf. also Quirk et al. 1972, p. 672).

*The ship arrived but the passengers could not get off.
The yacht is cheap but elegant.*

2. Comparison. In this function, the first conjunct is not so directly contradicted by the second. A contradiction exists, but we may have to

go through additional levels of implication to find it. Consider the sentence:

That basketball player is short, but he's very quick.

3. **Affirmation.** This use of “but” always follows a negative clause, and actually augments the meaning of the preceding clause by adding supporting information:

The disease not only killed thousands of people, but also ended a period of economic welfare.

In this section we consider only the first, or adversative, function of the coordinating conjunction “but.”

6.2.1 The Semantic Function of “But”. “But” introduces an element of surprise into discourse. Because it expresses some kind of contradiction, “but” has no role in the propositional calculus equivalent to the roles filled by “and” and “or.” Although there are logical formation rules using the conjunction operator (“and”) and the disjunction operator (“or”), there is no “but” operator. What, then, is the semantic role of “but”? We believe that its function should be described at the metalevel as one of many rules guiding the construction of partial theories. This is expressed below.

Metarule (BUT)

The formulas Φ **but** Ψ , Φ' **but** Ψ' , ... of a (formal representation of) paragraph P are to be interpreted as follows:

In the construction of any $T \in PT(\vec{P})$ instead of taking \oplus to be the union \cup of $\sigma \rightarrow T_\sigma$, take the union of $\sigma \rightarrow T_\sigma / \{\Psi, \Psi', \dots\}$.

The symbol $\sigma \rightarrow T_\sigma / \{\Psi, \Psi', \dots\}$ denotes a maximal consistent with $\{\sigma, \Psi, \Psi', \dots\}$ sub-theory of $\sigma \rightarrow T_\sigma$, and in general T/T' will be a maximal consistent with T' subtheory of T .

“But” is then an order to delete from background information everything contradicting Ψ , but to use what remains. Notice that “and” does not have this meaning; a model for Φ **and** Ψ will not contain any part of a theory that contradicts either of the clauses Φ or Ψ .

Typically this rule will be used to override defaults, to say that the expected consequences of the first conjunct hold except for the fact expressed by the second conjunct; for instance: *We were coming to see you, but it rained (so we didn't)*. The rule BUT is supposed to capture the “contrary-to-expectation” function of “but.”

We present now a simple example of building a model of a one-sentence paragraph containing “but.” We will use this example to explain how the rule BUT can be used. Using background information presented below, we will construct a partial model for this one-sentence paragraph.

Example 4

This yacht is cheap, but it is elegant.

Referential level (a fragment)

$cheap(x) \rightarrow \{\neg elegant(x); poor_quality(x); \neg expensive\}$	(c1)
$expensive(x) \leftrightarrow \neg cheap(x)$	(enc1)
$yacht(x) \rightarrow \{ship(x) \ \& \ small(x)\}$	(y1)
$elegant(x) \rightarrow \{\neg cheap(x), \dots\}$	(e1)
$elegant(x) \ \& \ yacht(x) \rightarrow \{status_symbol(x)\}$	(e_y1)

Note: Compare (y1) with (c1); in (y1) smallness is a property of a ship; this would be more precisely expressed as $yacht(x) \rightarrow [ship(x); \mathbf{property: small}(x)]$. This trick allows us not to conclude that “a big ant is big,” or “a small elephant is small.”

We ignore the problem of multiple meanings (theories) of predicates, and assume the trivial ordering in which all formulas are equally preferred. (But note that (e_y1) is still preferred to (e1) as a more specific theory of “elegant;” cf. Section 3.)

Construction of the model: In our case $\Phi \equiv yacht(y_0) \ \& \ cheap(y_0)$ and $\Psi \equiv elegant(y_0)$. Then

$$T_{cheap} = \{\neg elegant(x); poor_quality(x) ; \neg expensive(x)\}$$

$$T_{yacht} = \{ship(x) \ \& \ small(x)\}$$

$$T_{elegant} = \{expensive(x)\}$$

$$T_{yacht \ \& \ elegant} = \{status_symbol(x)\}$$

$$T_{cheap/\Psi} = \{poor_quality(x) ; \neg expensive(x)\}$$

$$T_{yacht/\Psi} = T_{yacht}$$

$$T_{yacht\&elegant/\Psi} = T_{yacht\&elegant}$$

We can now use the Metarule (BUT) and construct the partial theories of the sentence. In this case, there is only one:

$$PT(\Phi \text{ but } \Psi) = \{T\}, \text{ where}$$

$$T = \{yacht(y_0), \ elegant(y_0), \ cheap(y_0), \ ship(y_0) \ \& \ small(y_0), \\ status_symbol(y_0), \ poor_quality(y_0)\}$$

In other words, the yacht in question is a poor quality small and elegant ship serving as an inexpensive status symbol.

The partial model of the theory T is quite trivial: it consists of one entity representing the yacht and of a bunch of its attributes. However, the size of the model is not important here; what counts is the method of derivation of the partial theory.

6.2.2 Confirming the Analysis. The Metarule (BUT) is supposed to capture the “contrary-to-expectation” function of “but.” The next two rules follow from our formaliza-

tion; their validity indirectly supports the plausibility of our analysis of “but.”

BUT.C1 : Φ **but** $\neg\Psi$ is incorrect, if $\Phi \rightarrow \Psi$ is a “law.”

e.g. *Henry was murdered but not killed.*

Our referential level is a collection of partially ordered theories; we have expressed the fact that a theory of some ϕ is a “law” is by deleting the empty interpretation of ϕ from the partial order. If we accept the definition of a concept as given by necessary and sufficient conditions, the theories would all appear as laws. If we subscribe to a more realistic view where definitions are given by a collection of central/prototypical and peripheral conditions, only the peripheral ones can be contradicted by “but.” In either formalization we get BUT.C1 as a consequence: Since “laws” cannot be deleted, BUT can’t be applied, and hence its use in those kinds of sentences would be incorrect.

W. Labov (1973) discussed sentences of the form

**This is a chair but you can sit on it.*

The sentence is incorrect, since the function “one can sit on it” belongs to the core of the concept “chair”; so—contrary to the role of “but”—the sentence does not contain any surprising new elements. Using the Metarule (BUT) and the cooperative principle of Grice, we get

BUT.C2: Φ **but** Ψ is incorrect, if $\Phi \rightarrow \Psi$ is a “law.”

The Metarule (BUT) gives the semantics of “but;” the rules BUT.C1 and BUT.C2 follow from it (after formalization in a sufficiently strong metalanguage such as type theory or set theory). We can link all of them to procedures for constructing and evaluating models of text. Are they sufficient? Certainly not. We have not dealt with the other usages of “but;” neither have we shown how to deal with the apparent asymmetry of conclusions: *cheap but elegant* seems to imply “worth buying,” but *elegant but cheap* doesn’t; we have ignored possible prototypical effects in our semantics. However, we do believe that other rules, dealing with “but” or with other connectives, can be conveniently expressed in our framework. (The main idea is that one should talk explicitly and formally about relations between text and background knowledge, and that this knowledge is more than just a list of facts—it has structure, and it is ambiguous.) Furthermore, the semantics of “but” as described above is computationally tractable.

We also believe that one could not present a similarly natural account of the semantics of “but” in traditional logics, because classical logics withstand contradictions with great difficulty. Contradiction, however, is precisely what “but” expresses. Notice that certain types of coordinating conjunctions often have their counterparts in classical logic: copulative (*and, or, neither–nor, besides, sometimes* etc.), disjunctive (like *either–or*), illative (*hence, for that reason*). Others, such as explanatory (*namely* or *viz.*) or causal (*for*) conjunctions can probably be expressed somehow, for better or worse, within a classical framework. Thus the class of adversative conjunctions (*but, still, and yet, nevertheless*) is in that sense unique.

7. Conclusions

We hope that the reader has found this paper coherent, and its topic—the correspondence between paragraphs and models—interesting. Our strategy was to divide the

subject into three subtopics: a **theory of anaphora**, which corresponds to the logical *theory of equality* in p-models; a **theory of background knowledge**, expressed as the logical *theory of reference* in the three-level semantics; and **principles of communication** encoded in metarules. These principles include Gricean maxims and the semantics of cohesion, and specify a *model theory* for the three-level semantics. The framework resulting from putting these theories together is computational, empirical, and verifiable (even if incomplete); furthermore, it has strong links to already existing natural language processing systems. In particular, the new logical level—the referential level—is exemplified by on-line dictionaries and other reference works, from which we extract background information about defaults and plausibility rankings.

We also hope that the reader would be inclined to share our belief that natural language text can be properly and usefully analyzed by means of a three-level semantics that includes an object level, a metalevel, and a referential level. We believe that the coherence of an essay, a paper, or a book can be described by an extension of our theory. The work of van Dijk and Kintch (1983) on “macrostructures” could probably form the basis for such an expansion. Similarly, much of the abovementioned work by Hobbs, Webber, Grosz, and Sidner can be incorporated into this framework.

Our main intention was to demonstrate that formal notions of background knowledge can be used to

- define coherence, make it semantically distinct from mere consistency, and link it formally with the notion of a topic;
- define a class of p-models—logical models of paragraphs;
- provide a semantics for “but” (which exemplifies our understanding of grammatical cohesion);
- express the Gricean maxims formally, and use them in a computational model of communication (which seems to contradict Allen 1988, p. 464).

Moreover, we tried to convince the reader that paragraph is an important linguistic unit, not only because of its pragmatic importance exemplified by coherence and links to background knowledge, but also because of its role in assignment of syntactic structures (viz. ellipsis) and in semantics (viz. its possible role in evaluating semantic representations).

A great many issues have been omitted from our analysis. Thus, although we are aware that anaphora resolution and consistency depend on previously processed text, the problem of connecting a paragraph to such text has been conveniently ignored. Notice that this doesn’t make our thesis about paragraphs being units of semantic processing any weaker, we have not claimed that paragraphs are independent. The questions of how to translate from natural language to a logical notation needs a lot of attention; we have merely assumed that this can be done. Continuing this list, we have accepted a very classical theory of meaning, given by Tarski: the truth is what is satisfied in a model. This theory should be refined, for instance by formalizing Lakoff’s (1987) concept of radial categories, and proposing mechanisms for exploiting it. By the same token, the concept of reference has to be broadened to include iconic (e.g., visual and tactile) information. And certainly it would be nice to have a more detailed theory describing the role of the metalevel. In particular, we can imagine that the simple structure of a collection of set theoretic formulas can be replaced by something more interesting. We leave this as another open problem.

We have shown that it is possible to develop a formal system with an explicit relationship between background knowledge and text, showing mechanisms that take

advantage of preference, coherence, and contradiction (the reality of these phenomena has never been disputed, but their semantic functions had not been investigated). We should also mention that we have also begun some work on actually checking the empirical validity of this model (cf. e.g. Braden-Harder and Zadrozny 1990), using on-line dictionaries (LDOCE and Webster's) as the referential level. We know, of course, that existing dictionaries are very imperfect, but (1) they can be accessed and used by computer programs; (2) they are getting better, as they are very systematically created with help of computers (see Sinclair 1987 for an account of how COBUILD was constructed); (3) obviously, we can imagine useful new ones, like a Dictionary of Pragmatics; (4) we believe that we can explore the new inference mechanisms even in such unrefined environments.

Although the scheme we have proposed certainly needs further refinement, we believe that it is correct in two of its most important aspects: first, in the separation of current paragraph analysis (the object theory) from background information (the referential level); and second, in asserting that the function of a paragraph is to allow the building of a model of the situation described in the paragraph. This model can be stored, maybe modified, and subsequently used as a reference for processing following paragraphs.

Finally, the paper can be viewed as an argument that the meaning of a sentence cannot be defined outside of a context, just as the truth value of a formula cannot be computed in a vacuum. A paragraph is the smallest example of such a context—it is “a unit of thought.”

Acknowledgments

We gratefully acknowledge the help provided by our anonymous but diligent referees, by Graeme Hirst and Susan McRoy, and by our many colleagues at IBM Research, all for whom helped us to avoid worst consequences of Murphy's Law during the writing of this paper.

References

- Allen, J. (1987). *Natural Language Understanding*. Benjamin/Cummings. Menlo Park, California.
- Berwick, R.C. (1986). “Learning from Positive-Only Examples: The Subset Principle and Three Case Studies.” In Michalski, R.S. et al., eds. *Machine Learning Vol. II*. Morgan Kaufmann. Los Altos, California: 625–645.
- Black, J.B., and Bower, G.H. (1979). “Episodes as Chunks in Narrative Memory.” *Journal of Verbal Learning and Verbal Behavior* 18: 309–318.
- Bond, S.J., and Hayes, J.R. (1983). Cues People Use to Paragraph Text. *manuscript*. Dept. of Psychology, Carnegie-Mellon University.
- Brachman, R.J., Fikes, R.E., and Levesque, H.J. (1985). “Krypton: A Functional Approach to Knowledge Representation.” In Brachman, R.J., and Levesque, H.J., eds. *Readings in Knowledge Representation*. Morgan Kaufmann. Los Altos, California: 411–430.
- Braden-Harder, L., and Zadrozny, W. (1989). *Lexicons for Broad Coverage Semantics*. IBM Research Division Report, RC 15568.
- Chafe, W.L. (1979). “The Flow of Thought and the Flow of Language.” In Givón, T., ed., *Syntax and Semantics*, Vol. 12. Academic Press. New York, New York.
- Charniak, E. (1983). “Passing Markers: A Theory of Contextual Influence in Language Comprehension.” *Cognitive Science*, 7(3): 171–190.
- Charniak, E., and McDermott, D. (1985). *Introduction to Artificial Intelligence*. Addison-Wesley. Reading, Massachusetts.
- Crothers, E.J. (1979). *Paragraph Structure Inference*. Ablex Publishing Corp., Norwood, New Jersey.
- Etherington, D.W., and Mercer, R.E. (1987). “Domain Circumscription: A Reevaluation.” *Computational Intelligence* 3: 94–99.
- Frisch, A. (1987). “Inference without Chaining.” *Proc. IJCAI-87*. Milan, Italy: 515–519.
- Fowler, H.W. (1965). *A Dictionary of Modern English Usage*. Oxford University Press. New York, New York.
- Genesereth, M.R., and Nilsson, N.J. (1987). *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann. Los Altos, California.
- Grice, H.P. (1978). “Further Notes on Logic

- and Conversation." In Cole, P., ed. *Syntax and Semantics 9: Pragmatics*. Academic Press. New York, New York: 113–128.
- Grice, H.P. (1975). "Logic and Conversation." In Cole, P., and Morgan, J.L., eds., *Syntax and Semantics 3: Speech Acts*. Academic Press. New York, New York: 41–58.
- Groesser, A.C. (1981). *Prose Comprehension Beyond the Word*. Springer. New York, New York.
- Grosz, B.J. (1978). DISCOURSE KNOWLEDGE—Section 4 of Walker, D.E., ed., *Understanding Spoken Language*. North-Holland. New York, New York: 229–344.
- Grosz, B.J. (1977). "The Representation and Use of Focus in a System for Understanding Dialogs." *Proc. IJCAI-77*. W. Kaufmann. Los Altos, California: 67–76.
- Haberlandt, K., Berian, C., and Sandson, J. (1980). "The Episode Schema in Story Processing." *Journal of Verbal Learning and Verbal Behavior* 19: 635–650.
- Halliday, M.A.K., and Hasan, R. (1976). *Cohesion in English*. Longman Group Ltd. London.
- Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*. MIT Press. Cambridge, Massachusetts.
- Hinds, J. (1979). "Organizational Patterns in Discourse." In Givon, T., ed., *Syntax and Semantics*, Vol. 12. Academic Press. New York, New York.
- Hintikka, J. (1985). *The Game of Language*. D. Reidel. Dordrecht.
- Hirst, G. (1987). *Semantic Interpretation and the Resolution of Ambiguity*. Cambridge University Press. Cambridge.
- Hobbs, J.R. (1982). "Towards an Understanding of Coherence in Discourse." In Lehnert, W.G., and Ringle, M.H., eds. *Strategies for Natural Language Processing*. Lawrence Erlbaum. Hillsdale, New Jersey: 223–244.
- Hobbs, J.R. (1979). "Coherence and Coreference." *Cognitive Science* 3: 67–90.
- Hobbs, J.R. (1978). "Resolving Pronoun References." *Lingua* 44: 311–338.
- Hobbs, J.R. (1977). 38 Examples of Elusive Antecedents from Published Texts. Research Report 77-2. Department of Computer Science, City College, CUNY, New York.
- Hobbs, J.R. (1976). Pronoun Resolution. Research Report 76-1. Dept. of Computer Science. City College, CUNY, New York.
- Hobbs, J., Stickel, M., Martin, P., and Edwards, D. (1988). "Interpretation as Abduction." In *Proc. of 26th Annual Meeting of the Association for Computational Linguistics*, ACL: 95–103.
- Jackendoff, R. (1983). *Semantics and Cognition*. The MIT Press. Cambridge, Massachusetts.
- Jensen, K. (1988). "Issues in Parsing." In A. Blaser, ed., *Natural Language at the Computer*. Springer-Verlag, Berlin, Germany: 65–83.
- Jensen, K. (1986). Parsing Strategies in a Broad-Coverage Grammar of English. Research Report RC 12147. IBM T.J. Watson Research Center, Yorktown Heights, New York.
- Jensen, K., and Binot, J.-L. (1988). "Disambiguating Prepositional Phrase Attachments by Using On-line Dictionary Definitions." *Computational Linguistics* 13.3-4.251–260 (special issue on the lexicon).
- Johnson-Laird, P.N. (1983). *Mental Models*. Harvard University Press. Cambridge, Massachusetts.
- Kamp, H. (1981). "A Theory of Truth and Semantic Representation." In Groenendijk, J.A.G., et al. eds., *Formal Methods in the Study of Language*, I. Mathematisch Centrum, Amsterdam: 277–322.
- Labov, W. (1973). "The Boundaries of Words and Their Meanings." In Fishman, J., *New Ways of Analyzing Variation in English*. Georgetown U. Press. Washington, D.C.: 340–373.
- Lakoff, G. (1987). *Women, Fire and Dangerous Things*. The University of Chicago Press. Chicago, Illinois.
- Levesque, H.J. (1984). "A Logic of Implicit and Explicit Beliefs." *Proc. AAAI-84*. AAAI: 198–202.
- Longacre, R.E. (1979). The Paragraph as a Grammatical Unit. In Givon, T., ed., *Syntax and Semantics*, Vol. 12. Academic Press. New York, New York.
- Longman Dictionary of Contemporary English*. (1978). Longman Group Ltd., London.
- Lyons, J. (1969). *Introduction to Theoretical Linguistics*. Cambridge University Press. Cambridge, England.
- Mann, W.C., and Thompson, S.A. (1983). Relational Propositions in Discourse. Information Sciences Institute Research Report 83–115.
- Moens, M., and Steedman, M. (1987). "Temporal Ontology in Natural Language." *Proc. 25th Annual Meeting of the ACL*. Stanford, California: 1–7.
- Montague, R. (1970). "Universal Grammar." *Theoria* 36: 373–398.
- Mycielski, J. (1981). "Analysis without Actual Infinity." *Journal of Symbolic Logic*, 46(3): 625–633.

- Patel-Schneider, P.S. (1985). "A Decidable First Order Logic for Knowledge Representation." *Proc. IJCAI-85*. AAAI: 455–458.
- Perlis, D. (1985). "Languages with Self Reference I: Foundations." *Artificial Intelligence* 25: 301–322.
- Poole, D. (1988). "A Logical Framework for Default Reasoning." *Artificial Intelligence* 36(1): 27–47.
- Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1972). *A Grammar of Contemporary English*. Longman Group Ltd. London.
- Reggia, J.A. (1985). "Abductive Inference." In Karma, K.N., ed., *Expert Systems in Government Symposium*. IEEE: 484–489.
- Reiter, R. (1987). "A Theory of Diagnosis from First Principles." *Artificial Intelligence*, 32(1): 57–95.
- Sidner, C. (1983). "Focusing in the Comprehension of Definite Anaphora." In Brady, M., and Berwick, R., eds., *Computational Models of Discourse*, MIT Press. Cambridge, Massachusetts: 330–367.
- Sells, P. (1985). *Lectures on Contemporary Syntactic Theories*. CSLI lecture notes; no. 3.
- Shoenfield, J.R. (1987). *Mathematical Logic*. Addison-Wesley. New York, New York.
- Sinclair, J.M., ed. (1987). *Looking Up. An account of the COBUILD project*. Collins ELT, London.
- Small, S.L., Cottrell, G.W., and Tanenhaus, M.K., eds. (1988). *Lexical Ambiguity Resolution*. Morgan Kaufmann. San Mateo, California.
- Turner, M. (1987). *Death is the Mother of Beauty*. The University of Chicago Press. Chicago, Illinois.
- van Dijk, T.A., and Kintch, W. (1983). *Strategies of Discourse Comprehension*. Academic Press. Orlando, Florida.
- Warriner, J.E. (1963). *English Grammar and Composition*. Harcourt, Brace & World, Inc., New York, New York.
- Webber, B. (1983). "So What Can We Talk About Now?" In Brady, M., and Berwick, R., eds., *Computational Models of Discourse*. MIT Press. Cambridge, Massachusetts: 147–154.
- Webber, B. (1987). "The Interpretation of Tense in Discourse." *Proc. 25th Annual Meeting of the ACL*, ACL: 147–154.
- Webster's Seventh New Collegiate Dictionary*. (1963). Merriam-Webster, Inc. Springfield, Massachusetts.
- Woods, W. (1987). "Don't Blame the Tool." *Computational Intelligence* 3(3): 228–237.
- Zadrozny, W. (1987a). "Intended Models, Circumscription and Commonsense Reasoning." *Proc. IJCAI-87*: 909–916.
- Zadrozny, W. (1987b). "A Theory of Default Reasoning." *Proc. AAAI-87*. Seattle, Washington: 385–390.
- Zadrozny, W. (unpublished). *The Logic of Abduction*.

