Knowledge Representation for Commonsense Reasoning with Text

Kathleen Dahlgren
Joyce McDowell
IBM Los Angeles Scientific Center

Edward P. Stabler, Jr.[1]
University of Western Ontario

1 Introduction

1.1 Naive Semantics

The reader of a text actively constructs a rich picture of the objects, events, and situation described. The text is a vague, insufficient, and ambiguous indicator of the world that the writer intends to depict. The reader draws upon world knowledge to disambiguate and clarify the text, selecting the most plausible interpretation from among the (infinitely) many possible ones. In principle, any world knowledge whatsoever in the reader's mind can affect the choice of an interpretation. Is there a level of knowledge that is general and common to many speakers of a natural language? Can this level be the basis of an explanation of text interpretation? Can it be identified in a principled, projectable way? Can this level be represented for use in computational text understanding? We claim that there is such a level, called **naive semantics** (NS), which is commonsense knowledge associated with words. Naive semantics identifies words with concepts, which vary in type. Nominal concepts are categorizations of objects based upon naive theories concerning the nature and typical description of conceptualized objects. Verbal concepts are naive theories of the implications of conceptualized events and states.[2] Concepts are considered naive because they are not always objectively true, and bear only a distant relation to scientific theories. An informal example of a naive nominal concept is the following description of the typical lawyer.

1. If someone is a lawyer, typically they are male or female, well-dressed, use paper, books, and briefcases in their job, have a high income and high status. They are well-educated, clever, articulate, and knowledgeable, as well as contentious, aggressive, and ambitious. Inherently lawyers are adults, have gone to law school, and have passed the bar. They practice law, argue cases, advise clients, and represent them in court. Conversely, if someone has these features, he/she probably is a lawyer.

In the classical approach to word meaning, the aim is to find a set of primitives that is much smaller than the set of words in a language and whose elements can be conjoined in representations that are truth-conditionally adequate. In such theories "bachelor" is represented as a conjunction of primitive predicates.

2. bachelor(X) ⇔ adult(X) & human(X) & male(X) & unmarried(X)

In such theories, a sentence such as (3) can be given truth conditions based upon the meaning representation of "bachelor," plus rules of compositional semantics that map the sentence into a logical formula that asserts that the individual denoted by "John" is in the set of objects denoted by "bachelor."

3. John is a bachelor.

The sentence is true just in case all of the properties in the meaning representation of "bachelor" (2) are true of "John." This is essentially the approach in many computational knowledge representation schemes such as KRYPTON (Brachman et al. 1985), approaches following Schank (Schank and Abelson 1977), and linguistic semantic theories such as Katz (1972) and Jackendoff (1985).

Smith and Medin (1981), Dahlgren (1988a), Johnson-Laird (1983), and Lakoff (1987) argue in detail that all of these approaches are essentially similar in this way and all suffer from the same defects, which we summarize

0362-613X/ 89 /010149–170$03.00

briefly here. Word meanings are not scientific theories and do not provide criteria for membership in the categories they name (Putnam 1975). Concepts are vague, and the categories they name are sometimes vaguely defined (Labov 1973; Rosch et al. 1976). Membership of objects in categories is gradient, while the classical approach would predict that all members share full and equal status (Rosch and Mervis 1975). Not all categories can be decomposed into primitives (e.g., color terms). Exceptions to features in word meanings are common (most birds fly, but not all) (Fahlman 1979). Some terms are not intended to be used truth-conditionally (Dahlgren 1988a). Word meanings shift in unpredictable ways based upon changes in the social and physical environment (Dahlgren 1985b). The classical theory also predicts that fundamentally new concepts are impossible.

NS sees lexical meanings as naive theories and denies that meaning representations provide truth conditions for sentences in which they are used. NS accounts for the success of natural language communication, given the vagueness and inaccuracy of word meanings, by the fact that natural language is anchored in the real world. There are some real, stable classes of objects that nouns are used to refer to, and mental representations of their characteristics are close enough to true, enough of the time, to make reference using nouns possible (Boyd 1986). Similarly, there are real classes of events which verbs report, and mental representations of their implications are approximately true. The vagueness and inaccuracy of mental representations requires non-monotonic reasoning in drawing inferences based upon them. Anchoring is the main explanation of referential success, and the use of words for imaginary objects is derivative and secondary.

NS differs from approaches that employ exhaustive decompositions into primitive concepts which are supposed to be true of all and only the members of the set denoted by *lawyer*. NS descriptions are seen as heuristics. Features associated with a concept can be overridden or corrected by new information in specific cases (Reiter 1980). NS accounts for the fact that while English speakers believe that an inherent function of a lawyer is to practice law, they are also willing to be told that some lawyer does not practice law. A non-practicing lawyer is still a lawyer. The goal in NS is not to find the minimum set of primitives required to distinguish concepts from each other, but rather, to represent a portion of the naive theory that constitutes the cognitive concept associated with a word. NS descriptions include features found in alternative approaches, but more as well. The content of features is seen as essentially limitless and is drawn from psycholinguistic studies of concepts. Thus, in NS, featural descriptions associated with words have as values not primitives, but other words, as in Schubert et al. (1979).

In NS, the architecture of cognition that is assumed is one in which syntax, compositional semantics, and

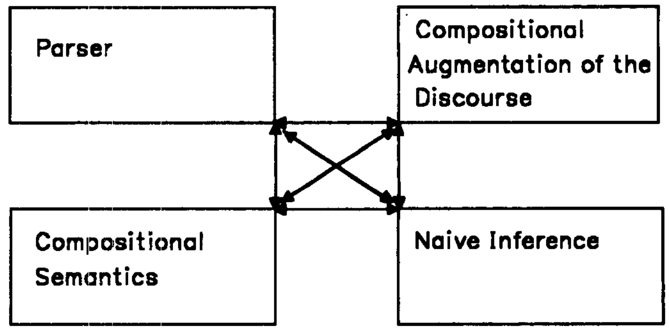

Figure 1. Components of Grammar in NS.

naive semantics are separate components with unique representational forms and processing mechanisms. Figure 1 illustrates the components. The autonomous syntactic component draws upon naive semantic information for problems such as prepositional phrase attachment and word sense disambiguation. Another autonomous component interprets the compositional semantics and builds **discourse representation structures** (DRSs) as in Kamp (1981) and Asher (1987). Another component models naive semantics and completes the discourse representation that includes the implications of the text. All of these components operate in parallel and have access to each other's representations whenever necessary.

1.2 THE KT SYSTEM

Naive semantics is the theoretical motivation for the KT system under development at the IBM Los Angeles Scientific Center by Dahlgren, McDowell, and others.[3] The heart of the system is a commonsense knowledge base with two components, a commonsense ontology and databases of generic knowledge associated with lexical items. The first phase of the project, which is nearly complete, is a text understanding and query system. In this phase, text is read into the system and parsed by the MODL parser (McCord 1987), which has very wide coverage. The parse is submitted to a module (DISAMBIG) that outputs a logical structure which reflects the scope properties of operators and quantifiers, correct attachment of post-verbal adjuncts, and selects word senses. This is passed to a semantic translator whose output (a DRS) is then converted to first-order logic (FOL). We then have the text in two different semantic forms (DRS and FOL), each of which has its advantages and each of which is utilized in the system in different ways. Queries are handled in the same way as text. Answers to the queries are obtained either by matching to the FOL textual database or to the commonsense databases. However, the commonsense knowledge is accessed at many other stages in the processing of text and queries, namely in parse disambiguation, in lexical retrieval, anaphora resolution, and in the construction of the discourse structure of the entire text.

The second phase of the project, which is at present in the research stage, will be to use the commonsense knowledge representations and the textual database to guide text selection. We anticipate a system, NewSelector, which, given a set of user profiles, will distribute textual material in accordance with user interests, thus, in effect, acting as an automatic clipping service. Our target text is the *Wall Street Journal*. The inferencing capabilities provided by the commonsense knowledge will allow us to go well beyond simple keyword search.

The theoretical underpinnings and practical work on the KT system have been reported extensively elsewhere, in conference papers (Dahlgren and McDowell 1986a, 1986b; McDowell and Dahlgren 1987) and in a book (Dahlgren 1988a). Since the publication of those works, a number of significant additions and modifications have been made to the system. The intended focus of this paper is this new work. However, in order to make this accessible to readers unfamiliar with our previous reports, we present in Section 2 an overview of the components of the present system. (Readers familiar with our system can skip Section 2). In the remaining sections on new work we have emphasized implementation, because this paper is addressed to the computational linguistics community: in Section 3, the details of disambiguation procedures that use the NS representations and in Section 4 the details of the query system. Finally, Section 5 discusses work in progress regarding discourse and naive semantics.

2 Overview of the KT System

2.1 Knowledge Representation

2.1.1 The Ontology

Naive theories associated with words include beliefs concerning the structure of the actual world and the significant "joints" in that structure. People have the environment classified, and the classification scheme of a culture is reflected in its language. Since naive semantics is intended as a cognitive model, we constructed the naive semantic ontology empirically, rather than intuitively. We studied the behavior of hundreds of verbs and determined selectional restrictions, which are constraints reflecting the naive ontology embodied in English. We also took into account psychological studies of classification (Keil 1979), and philosophical studies of epistemology (Strawson 1953).

2.1.1.1 Mathematical Properties of the Ontology

The ontology has several properties which distinguish it from classical taxonomies. It is a directed acyclic graph, rather than a binary tree, because many concepts have more than two subordinate concepts (Rosch et al. 1976). FISH, BIRD, MAMMAL, and so on, are subordinates of VERTEBRATE. It is a directed graph rather than a tree, because it handles cross-classification. Cross-classification is justified by contrasts between individual and collective nouns such as "cow" and "herd." This

ENTITY	→	(ABSTRACT v REAL) & (INDIVIDUAL v COLLECTIVE)
ABSTRACT	→	IDEAL v PROPOSITIONAL v NUMERICAL v IRREAL
NUMERICAL	→	NUMBER v MEASURE
REAL	→	(PHYSICAL v TEMPORAL v SENTIENT) & (NATURAL v SOCIAL)
PHYSICAL	→	(STATIONARY v NONSTATIONARY) & (LIVING v NONLIVING)
NON-STATIONARY	→	(SELFMOVING v NONSELFMOVING)
COLLECTIVE	→	MASS v SET v STRUCTURE
TEMPORAL	→	RELATIONAL v NONRELATIONAL
RELATIONAL	→	(EVENT v STATIVE) & (MENTAL v EMOTIONAL v NONMENTAL)
EVENT	→	(GOAL v NONGOAL) & (ACTIVITY v ACCOMPLISHMENT v ACHIEVEMENT)

Table 1. The Ontological Schema

implies that cognitively there are essentially parallel ontological schemas for individuals and collectives. Thus we have the parallel ontology fragments in Figure 2. Table 2 illustrates cross-classification at the root of the ontology, where ENTITY cross-classifies as either REAL or ABSTRACT, and as either INDIVIDUAL or COLLECTIVE. Cross-classification is handled as in McCord (1985, 1987).

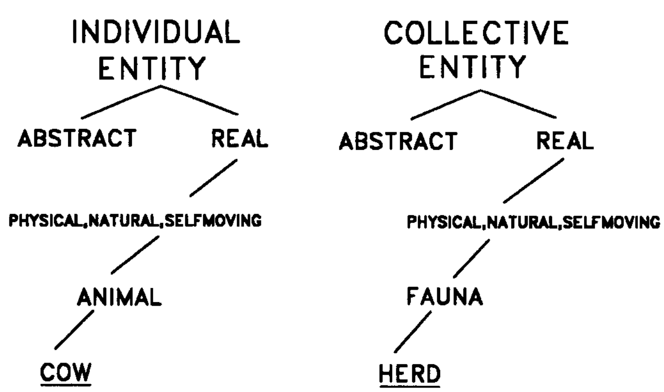

Figure 2. Parallel Portions of the Ontology.

	INDIVIDUAL	COLLECTIVE
REAL	*cow*	*herd*
ABSTRACT	*idea*	*book*

Table 2. Entity Node Cross Classification

Multiple attachments of instantiations to leaves is possible. For example, an entity, John, is both a HUMAN with the physical properties of a MAMMAL, and is also a PERSON, a SENTIENT. As a SENTIENT, John can be the subject of mental verbs such as *think* and *say*. Institutions are also SENTIENTs, so that the SENTIENT node reflects English usage in pairs like (4).

4. John sued Levine. The government sued Levine.

On the other hand, John, as a HUMAN, is like animals, and has physical properties. Both John and a cow can figure as subjects of verbs like "eat" and "weigh." Multiple attachment was justified by an examination of the texts. It was found that references to human beings in text, for example, deal with them either as persons (SENTIENTs) or as ANIMALs (physiological beings), but rarely as both at the same time.

2.1.1.2 ONTOLOGICAL CATEGORIES

The INDIVIDUAL/COLLECTIVE cut was made at the level of ENTITY (the highest level) because all types of entities are conceived individually or in collections. COLLECTIVE breaks into sets of identical members (herd, mob, crowd, fleet), masses that are conceived as stuff (sand, water), and structures where the members have specified relations, such as in institutions (school, company, village). Leaf node names, such as ANIMAL and FAUNA, are shorthand for collections of categories inherited from dominating nodes and by cross-classification. Thus "cow" and "herd" share all categories except that "cow" is an INDIVIDUAL term and "herd" is a COLLECTIVE term.

The REAL node breaks into the categories PHYSICAL, TEMPORAL, and SENTIENT, and also NATURAL and SOCIAL. Entities (or events) that come into being (or take place) naturally must be distinguished from those that arise through some sort of social intervention. Table 3 illustrates the assignment of example words under the REAL cross-classification.

	INDIVIDUAL		COLLECTIVE	
	NATURAL	SOCIAL	NATURAL	SOCIAL
PHYSICAL	*rock*	*knife*	*sand*	*fleet*
SENTIENT	*man*	*programmer*	*mob*	*clinic*
TEMPORAL	*earthquake*	*party*	*winter*	*epoch*

Table 3. Attachment of Nouns under REAL

The SENTIENT/PHYSICAL distinction is placed high because in commonsense reasoning, the properties of people and things are very different. Verbs select for SENTIENT arguments or PHYSICAL arguments, as illustrated in (5). Notice also, that as a physical object, an individual entity like John can be the subject both of verbs that require physical subjects and those that require SENTIENT subjects, as in (6).

5. The lawyer/the grand jury indicted Levine.
   *The cow indicted Levine.
6. John fell.
   John sued Levine.
   The cow fell.

We make the SENTIENT/NON-SENTIENT distinction high up in the hierarchy for several reasons. Philosophically, the most fundamental distinction in epistemology (human knowledge) is arguably that between thinking and non-thinking beings (Strawson 1953). Psychology has shown that infants are able to distinguish humans from all other objects and they develop a deeper and more complex understanding of humans than of other objects (Gelman and Spelke 1981). In the realm of linguistics, a class of verbs selects for persons, roles, and institutions as subjects or objects. Thus the SENTIENT distinction captures the similarity between persons and institutions or roles. There is a widespread lexical ambiguity between a locational and institutional reading of nouns, which can be accounted for by the SENTIENT distinction, as in (7).

7. The court is in the center of town.
   The court issued an injunction.

The NATURAL/SOCIAL distinction also was placed high in the hierarchy. Entities (including events) that are products of society, and thereby have a social function, are viewed as fundamentally different from natural entities in the commonsense conceptual scheme. The distinction is a basic one psychologically (Miller 1978; Gelman and Spelke 1981). SOCIAL entities are those that come into being only in a social or institutional setting, with "institution" being understood in the broadest sense, for instance family, government, education, warfare, organized religion, etc.

2.1.1.3 CONSTRUCTION OF THE ONTOLOGY

The ontological schema was constructed to handle the selectional restrictions of verbs in 4,000 words of geography text and 6,000 words of newspaper text. These were arranged in a hierarchical schema. The hierarchy was examined and modified to reflect cognitive, philosophical, and linguistic facts, as described above. It was pruned to make it as compact as possible. We minimized empty terminal nodes. A node could not be part of the ontology unless it systematically pervaded some subhierarchy. Distinctions found in various places were relegated to feature status. The full ontology with examples may be found in Dahlgren (1988).

2.1.1.4 VERBS

In KT, verbs are attached to the main ontology at the node TEMPORAL because information concerning the temporality of situations described in a sentence is encoded on the verb as tense and because the relations indicated by verbs must be interpreted with respect to their location in time in order to properly understand the discourse structure of a text. Thus the ontology implies that events are real entities, and that linguistic, not conceptual, structure distinguishes verbalized from nominalized versions of events and states.

We view the cognitive structure of the concepts associated with nouns and verbs as essentially different. Non-derived nouns in utterances refer to objects. Lexical nouns name classes of entities that share certain features. Verbs name classes of events and states, but these do not share featural descriptions. Psychologically, verbs are organized around goal orientation and argument types (Huttenlocher and Lui 1979; Graesser and Hopkinson 1987).

The primary category cut at the node TEMPORAL is between nouns, which name classes of entities, in this case temporal entities like "party," "hurricane," and "winter," and verbs, which indicate relations between members of these nominal classes, like "hit," "love," "remember." We attach temporal nouns to the TEMPORAL/NON-RELATIONAL node and verbs to the TEMPORAL/RELATIONAL node. Many nouns are, of course, relational, in the sense that "father" and "indictment" are relational. Our node RELATIONAL does not carry this intuitive sense of relational, but instead simply indicates that words attached here require arguments for complete interpretation. So, while "father" is relational, it is possible to use "father" in a text without mentioning the related entity. But verbs require arguments (usually overtly, but sometimes understood, as in the case of commands) for full interpretation. Nominalizations like "indictment" are a special case. In our system, all deverbal nominalizations are so marked with a pointer to the verb from which they are derived. Subsequent processing is then directed to the verb, which, of course, is attached under TEMPORAL, RELATIONAL.

2.1.1.5 THE VENDLER CLASSIFICATION

One basis of the relational ontology is the Vendler (1967) classification scheme, which categorizes verbs into aspectual classes (see Dowty 1979). According to this classification, RELATIONAL divides into EVENT or STATIVE, and EVENT divides into ACTIVITY, ACHIEVEMENT, or ACCOMPLISHMENT. Vendler and others (particularly Dowty 1979) have found the following properties, which distinguish these classes.

8. STATIVE and ACHIEVEMENT verbs may not appear in the progressive, but may appear in the simple present.
   ACTIVITY and ACCOMPLISHMENT verbs may appear in the progressive and if they appear in the simple present, they are interpreted as describing habitual or characteristic states.
   ACCOMPLISHMENTs and ACHIEVEMENTs entail a change of state associated with a terminus (a clear endpoint). STATIVEs ("know") and ACTIVITYs ("run") have no well-defined terminus.
   ACHIEVEMENTs are punctual (*John killed Mary*) while ACCOMPLISHMENTs are gradual (*John built a house*).
   STATEs and ACTIVITYs have the subinterval property (cf. Bennett and Partee 1978).

Table 4 summarizes these distinctions. There are several standard tests for the Vendler (1967) system which can be found in Dowty (1979) and others and which we apply in classification.

The Vendler classification scheme is actually more accurately a classification of verb phrases than verbs. KT handles this problem in two ways. First, in sentence processing we take into account the arguments in the verb phrase as well as the verb classification to determine clause aspect, which can be any of the Vendler classes. Second, we classify each sense of a verb separately.

	Activity	Accomplishment	Achievement	State
	run, think	*build a house, read a novel*	*recognize,find*	*have, want*
Progressives	+	+	−	−
Terminus	−	+	+	−
Change of State	−	Gradual	Punctual	−
Subinterval Property	+	−	−	+

Table 4. The Vendler Verb Classification Scheme

The other nodes in the relational ontology are motivated by the psycholinguistic studies noted above. The MENTAL/NONMENTAL/EMOTIONAL distinction is made at the highest level for the same reasons that led us to place SENTIENT at a high level in the main ontology. All EVENTs are also cross-classified as GOAL oriented or not. This is supported by virtually every experimental study on the way people view situations, i.e., GOAL orientation is the most salient property of events and actions. For example, Trabasso and Van den Broek (1985) find that events are best recalled which feature the goals of individuals and the consequences of goals and Trabasso and Sperry (1985) find that the salient features of events are goals, antecedents, consequences, implications, enablement, causality, motivation, and temporal succession and coexistence. This view is further supported by Abbott, Black, and Smith (1985) and Graesser and Clark (1985). NONGOAL, ACCOMPLISHMENT is a null category because ACCOMPLISHMENTS are associated with a terminus and thus inherently GOAL oriented. On the other hand NONGOAL, ACHIEVEMENT is not a null category because the activity leading up to an achievement is always totally distinct from the achievement itself. SOCIAL, NONGOAL, ACTIVITY is a sparse category.

Cross-classifications inherited from the TEMPORAL node are SOCIAL/NATURAL and INDIVIDUAL/COLLECTIVE. The INDIVIDUAL/COLLECTIVE distinction is problematical for verbs, because all events can be viewed as a collection of an infinitude of subevents.

2.1.2 GENERIC KNOWLEDGE

Generic descriptions of the nouns and verbs were drawn from psycholinguistic data to the extent possible. In a typical experiment, subjects are asked to freelist features "characteristic" of and common to objects in categories such as DOG, LEMON, and SECRETARY (Rosch et al. 1976; Ashcraft 1976; Dahlgren 1985a). The number of subjects in such an experiment ranges from 20 to 75. Any feature that is produced in a freelisting experiment by several subjects is likely to be shared in the relevant subpopulation (Rosch 1975; Dahlgren 1985a). Features that were freelisted by at least one-fifth of the subjects are chosen for a second experiment, in which different subjects are asked to rate the features for typicality. Those features rated as highly typical by the second group can be considered a good first approximation to the content of the cognitive structures associated with the terms under consideration. The number of features shared in this way for a term averaged 15.

The generic knowledge in the KT system is contained in two generic data bases, one for nouns and one for verbs. There is a separate entry for each sense of each lexical item. The content of the entries is a pair of lists of features drawn from the psycholinguistic data (as described above) or constructed using these data as a model. A feature is, informally, any bit of knowledge that is associated with a term. In informal terms, these can be any items like "wears glasses" (programmer), "is red" (brick), or "can't be trusted" (used-car salesman). For each entry, the features are divided into two lists, one for typical features and one for inherent features. For example, a brick is typically red, but blood is inherently red. Together the lists comprise the entry description of the term that heads the entry.

The source of descriptions of social roles were data collected by Dahlgren (1985a). For physical objects we used generic descriptions from Ashcraft (1976), including raw data generously supplied by the author. An informal conceptualization for "lawyer" was shown in (1). The corresponding generic description is shown in (9). Features of the same feature type within either the inherent or typical list are AND'ed or OR'ed as required. Some features contain first-order formulas like the conditions in discourse representations. For example, one function feature has (advise(E,noun,Y) & client(Y) & regarding(E,Z) & law(Z)). The first argument of the predicate advise is an event reference marker. This event is modified in the regarding predicate. The second argument of advise is instantiated in the processing as the same entity that is predicated as being in the extension of the noun generically described in the representation, in this case, "lawyer."

```
9. lawyer(
   {behavior(contentious),appearance(well-dressed),
   status(high),income(high),
   sex(male),sex(female),tools(paper),tools(books),
     tools(briefcase),
   function(negotiate(*,noun,Y) & settlement(Y)),
   internal_trait(ambitious),internal_trait(articulate),
   internal_trait(aggressive),
   internal_trait(clever),
   internal_trait(knowledgeable)},
   {age(adult),education(law_school),
   legal_req(pass(*,noun,X) & bar(X))
   function(practice(*,noun,Y) & law(Y)),
   function(advise(E,noun,Y) & client(Y) &
     regarding(E,Z) & law(Z)),
   function(represent(E,noun,Y) & client(Y) &
   in(E,Z) & court(Z)),
   function(argue(*,noun,Y) & case(Y))}).
```

The entire set of features for nouns collected in this way so far sort into 38 feature types. These are age, agent, appearance, association, behavior, color, construction, content, direction, duration, education, exemplar, experienced-as, in-extension-of, frequency, function, goal, habitat, haspart, hasrole, hierarchy, internal-trait, legal-requirement, length, level, location, manner, material, name, object, odor, operation, owner, partof, physiology, place, processing, propagation, prototype, relation, requirement, rolein, roles, sex, shape, size, source, speed, state, status, strength, structure, taste, texture, time. A given feature type can be used in either typical or inherent feature lists. Since these are not primitives, we expect the list of feature types to expand as we enlarge the semantic domain of the system.

There is a much smaller set of feature types for verbs. We were guided by recent findings in the psycholinguistic literature which show that the types of information that subjects associate with verbs are substantially different from what they associate with nouns. Huttenlocher and Lui (1979) and Abbott, Black, and Smith (1985) in particular have convincingly argued that subjects conceive of verbs in terms of whether or not the activities they describe are goal oriented, the causal and temporal properties of the events described, and the types of entities that can participate as arguments to the verb. For the actual feature types, we adapted the findings of Graesser and Clark (1985), whose research focused on the salient implications of events in narratives. A small number of feature types is sufficient to represent the most salient features of events. These can be thought of as answers to questions about the typical event described by the verb that heads the entry. In addition, selectional restrictions on the verb are also encoded as a feature type. Feature types for verbs are cause, goal, what_enabled, what_happened_next, consequence_of_event, where, when, implies, how, selectional_restriction. An example of a generic entry for verbs follows.

10.

buy(	If someone buys something,
{what_enabled(can(afford(subj,obj))),	typically he can afford it,
how(with(X) & money(X)),	he uses money,
where(in(Y) & store(Y)),	he buys it in a store,
cause(need,subj,obj)),	he needs it,
what_happened_next(use(subj,obj))},	and later he uses it.
{goal(own(subj,obj)),	Inherently, his goal is to own it,
consequence_of_event(own((subj,obj)),	and after buying it, he does own it.
selectional_restriction(sentient(subj)),	The buyer is sentient and
implies(merchandise(obj))}).[4]	what is bought is merchandise.

Naive Semantic representations of generic knowledge contain fifteen or more pieces of information per word, relatively more than required by other theories. The magnitude of the lexicon is counterbalanced by constraints that naturally obtain in the generic knowledge. Study of the protocols of subjects in prototype experiments reveals that people conceive of objects in constrained patterns of feature types. For example, animals are conceived in terms of physical and behavioral properties, while social roles are conceived in terms of functional and relational properties. Thus not all feature types occur in the representations of all nouns (or verbs). The pattern of features relevant to each node in the ontology is called a **Kind Type**. Each feature is classified by type as a COLOR, SIZE, FUNCTION, INTERNAL TRAIT or other. At each node, only certain feature types are applicable. Features at lower nodes inherit feature type constraints from the nodes above them in the ontology. For instance, any node under SOCIAL may have certain feature types, and any node under ROLE may have those feature types inherited from SOCIAL, as well as further feature types. Examples contrasting "elephant" and "lawyer" are shown in Tables 5 and 6.

Node in Ontology	Feature types associated with the node	Feature values for *elephant*
ENTITY	haspart	trunk
	haspart	4 legs
	partof	herd
PHYSICAL	color	grey
	size	vehiclesized
	texture	rough
LIVING	propagation	live births
	habitat	jungle
ANIMAL	sex	male or female
	behavior	lumbers
	behavior	eats grass

Table 5. Animal Kind Type

A lexical augmentation facility is used to create generic entries. This facility exploits the fact that possible feature types for any term are constrained by the ontological attachment of the term, by the Kind Type to which they belong (Dahlgren & McDowell 1986a; Dahlgren 1988a). For example, it is appropriate to encode the feature type "behavior" for "dog" but not for "truck." Similarly, it is appropriate to encode the feature type "goal" for "dig" but not for "fall." The lexical augmentation facility presents the user with appropriate choices for each term and then converts the entries to a form suitable for processing in the system.

2.2 TEXT INTERPRETATION ARCHITECTURE

2.2.1 PARSER

The overall goal of the KT project is text selection based on the extraction of discourse relations guided by

Node in Ontology	Feature types associated with the node	Feature values for *lawyer*
SOCIAL	function	types
	function	practice law
	function	argue cases
	function	advise clients
	function	represent clients in court
	requirement	pass bar
	appearance	well-dressed
SENTIENT	internaltrait	friendly
	education	law school
	internaltrait	clever
	internaltrait	articulate
	internaltrait	contentious
	sex	male
	sex	female
ROLE	relation	high status
	income	high
	tools	books
	tools	briefcases

Table 6. Social Role Kind Type

naive semantic representations. This goal motivated the choice of DRT as the compositional semantic formalism for the project. The particular implementation of DRT which we use assumes a simple, purely syntactic parse as input to the DRS construction procedures (Wada and Asher 1986). Purely syntactic parsing and formal semantics are unequal to the task of selecting one of the many possible parses and interpretations of a given text, but human readers easily choose just one interpretation. NS representations are very useful in guiding this choice. They can be used to disambiguate parse trees, word senses, and quantifier scope. We use an existing parser (MODL) to get one of the syntactic structures for a sentence and modify it based upon NS information. This allows us to isolate the power of NS representations with respect to the parsing problem. Not only are NS representations necessary for a robust parsing capability, but also in anaphora resolution and discourse reasoning. Furthermore, the parse tree must be available to the discourse coherence rules. Thus our research has shown that not only must the NS representations be accessible at all levels of text processing, but purely syntactic, semantic, and pragmatic information that has been accumulated must also be available to later stages of processing. As a result, the architecture of the system involves separate modules for syntax, semantics, discourse and naive semantics, but each of the modules has access to the output of all others, as shown in Figure 1.

The parser chosen is the **Modular Logic Grammar** (McCord 1987), (MODL). Both MODL and KT are written in VM/PROLOG (IBM 1985). The input to MODL is a sentence or group of sentences (a text). In KT we intercept the output of MODL at the stage of a labeled bracketing marked with grammatical features and before any disambiguation or semantic processing is done. In effect, we bypass the semantics of MODL in order to test our NS representations. In the labeled bracketing each lexical item is associated with an argument structure that can be exploited in semantic interpretation. The labeled bracketing output by MODL is slightly processed before being passed to our module DISAMBIG. Here the commonsense knowledge base is accessed to apply rules for prepositional phrase attachment (Dahlgren and McDowell 1986b) and word sense disambiguation (Dahlgren 1988a), as well as to assign the correct scope properties to operators and quantifiers. The output is a modified parse. All of these modules are in place and functional. The word sense disambiguation rules are in the process of being converted. An example of the input to DISAMBIG and the resulting output is as follows:

11. Input S:

John put the money in the bank.

Input to DISAMBIG:

```
s(np(n(n(john(V1)))) &
vp(v(fin(pers3,sg,past,ind),put(V1,V2)) &
np(detp(the(V3,V4)) & n(n(money(V2))) &
pp(p(in(V2,V5)) & np(detp(the(V6,V7)) &
  n(n(bank(V5))))))))
```

Output of DISAMBIG:

```
s(np(n(n(john(V1)))) &
vp(v(fin(pers3,sg,past,ind),put1(V1,V2)) &
np(detp(the(V3,V4)) & n(n(money(V2)))) &
pp(p(in(V2,V5)) & np(detp(the(V6,V7)) &
  n(n(bank2(V5)))))))
```

The differences are that the PP "in the bank" is VP-attached in the output of DISAMBIG rather than NP-attached as in the output of MODL, and that the words "put" and "bank" are assigned index numbers and changed to put1 and bank2, selecting the senses indicated by the word sense disambiguation algorithm.

2.2.2 SENTENCE-LEVEL SEMANTICS

The modified parse is then submitted to a semantics module, which outputs a structure motivated in part by current versions of discourse representation theory (DRT) (Kamp 1981; Wada and Asher 1986; Asher and Wada 1988). The actual form of the discourse representation structure (DRS) and its conditions list in the KT system differ from standard formats in DRT in that tense arguments have been added to every predicate and tense predicates link the tense arguments to the tense of the verb of the containing clause. The analysis of questions and negation was carried out entirely with respect to the KT system and to serve its needs. The DRT semantics is in place and functional for most structures covered by MODL. The commonsense knowledge representations are accessed in the DRT module for semantic interpretation of modals, the determination of sentence-internal pronoun anaphora (where simple C-command and agreement tests fail), and to determine some cases of quantifier scoping.

2.2.3 DISCOURSE-LEVEL SEMANTICS

As each sentence of a text is processed it is added to the DRS built for the previous sentence or sentences. Thus an augmented DRS is built for the entire text. In the augmentation module the commonsense knowledge representations are accessed to determine definite noun phrase anaphora, sentence-external pronoun anaphora, temporal relations between the tense predicates generated during sentence-level DRS construction, discourse relations (such as causal relations) between clauses, and the rhetorical structure of the discourse as a whole. The discourse work is being carried out mainly by Dahlgren and is in various stages of completion.

2.2.4 FOL

Since standard proof techniques are available for use with logical forms, the DRS formulated by the sentence-level and discourse-level semantic components is converted to standard logic. A number of difficulties present themselves here. In the first place, given any of the proposed semantics for DRSs (e.g., Kamp 1981; Asher 1987), DRSs are not truth functional. That is, the truth value of a DRS structure (in the actual world) is not generally a function of the truth values of its constituents. For example, this happens when verbs produce opaque contexts (Asher 1987). Since general proof methods for modal logics are computationally difficult, we have adopted the policy of mapping DRSs to naive first-order translations in two steps, providing a special and incomplete treatment of non-truth functional contexts. The first step produces representations that differ minimally from standard sentences of first-order logic. The availability of this level of representation enhances the modularity and the extensibility of the system. Since first-order reasoning is not feasible in an application like this, a second step converts the logical forms to clausal forms appropriate for the problem solver or the textual knowledge base. We describe each of these steps in turn.

2.2.4.1 THE TRANSLATION TO STANDARD LOGICAL FORMS

The sentence-level and discourse-level semantic components disambiguate names and definite NPs, so that each discourse reference marker is the unique canonical name for an individual or object mentioned in the discourse. The scoping of quantifiers and negations has also been determined by the semantic processing. This allows the transformation of a DRS to FOL to be rather simple. The basic ideas can be briefly introduced here; a more thorough discussion of the special treatment of queries is provided below. The conditions of a DRS are conjoined. Those conditions may include some that are already related by logical operators (if-then, or, not), in which case the logical form includes the same operators. Discourse referents introduced in the consequent of an if-then construction may introduce quantifiers: these are given narrow scope relative to quantifiers in the consequent (cf. Kamp's notion of a "subordinate"

DRS in Kamp 1981; Asher and Wada 1988). Any quantifiers needing wide scope will already have been moved out of the consequent by earlier semantic processing. The DRSs of questions contain special operators that, like the logical operators, take DRS arguments that represent the scope of the material being questioned. A similar indication is needed in the logical form. In the case of a yes-no question, we introduce a special vacuous quantifier just to mark the scope of the questioned material for special treatment by the problem solver (see below). In the case of wh-questions, a special wh-quantifier is introduced, again indicating the scope of the questioned material and triggering a special treatment by the problem solver. Verbs of propositional attitude and other structures with opaque contexts are treated as, in effect, introducing new isolated subtheories or "worlds." For example, "John believes all men are mortal" is represented as a relation of belief holding between John and a logical form (see below for more detail), though it is recognized that this approach will need to be supplemented to handle anaphora and propositional nominals (cf., e.g., Asher 1988).

2.2.4.2 The Conversion to Specialized Clausal Forms

Considerations of efficiency motivate a further transformation in our logical forms. After the first step of processing, we have standard first-order logical forms, except that they may include special quantifiers indicating questioned material. Consider first those logical forms that are not inside the scope of any question quantifier. These are taken as representations of potential new knowledge for the textual data base. Since the inference system must solve problems without user guidance, it would be infeasible to reason directly from the first-order formulations. Clausal forms provide an enormous computational advantage. For these reasons, we transform each sentence of first-order logic into a clausal form with a standard technique (cf., e.g., Chang and Lee 1973), introducing appropriate Skolem functions to replace existentially quantified variables. The textual database can then be accessed by a clausal theorem prover. In the current system, we use efficient Horn clause resolution techniques (see below), so the knowledge representation is further restricted to Horn clauses, since completeness is less important than feasible resource use in the present application. Definite clauses are represented in a standard Prolog format, while negative clauses are transformed into definite clauses by the addition of a special positive literal "false(n)," where n is an integer that occurs in no other literal with this predicate. This allows a specialized incomplete treatment of negation-as-inconsistency (cf. Gabbay and Sergot 1986). The definite clause translations of the text can then be inserted into a textual knowledge base for use by the reasoning component. The presence of question quantifiers triggers a special treatment in the conversion to clausal form. Our problem solver uses standard resolution techniques: to prove a proposition, we show that its negation is incompatible with the theory. Accordingly, the material inside the scope of a question operator is treated as if it were negated, and this implies an appropriately different treatment of any quantifiers inside the scope of the operators.

2.2.5 Reasoner

In the architecture of the system, the reasoning module is broken into two parts: the specialized query processing system and a general purpose problem solver. The special processing of queries is described in detail below. The problem solver is based on a straightforward depth-bounded Horn clause proof system, implemented by a Prolog metainterpreter (e.g. Sterling and Shapiro 1986). The depth bound can be kept fixed when it is known that no proofs should exceed a certain small depth. When a small depth bound is not known, the depth can be allowed to increase iteratively (cf. Stickel 1986), yielding a complete SLD resolution system. This proof system is augmented with negation-as-failure for predicates known to be complete (see the discussion of open and closed world assumptions below), and with a specialized incomplete negation-as-inconsistency that allows some negative answers to queries in cases where negation-as-failure cannot be used.

2.2.6 Relevance

The RELEVANCE module will have the responsibility of determining the relevance of a particular text to a particular user. The text and user profiles will be processed through the system in the usual way resulting in two textual data bases, one for the target text and one for the profile. Target and profile will then be compared for relevance and a decision made whether to dispatch the target to the profiled user or not. The relevance rules are a current research topic. The commonsense knowledge representations will form the primary basis for determining relevance.

3 Naive Semantics in the KT System

From the foregoing brief overview of the KT system, it should be clear that naive semantics is used throughout the system for a number of different processing tasks. In this section we show why each of these tasks is a problem area and how NS can be used to solve it.

3.1 Prepositional Phrase Attachment[5]

The proper attachment of post-verbal adjuncts is a notoriously difficult task. The problem for prepositional phrases can be illustrated by comparing the following sentences.

12. [S The government [VP had uncovered [NP an entire file [PP about the scheme]]]].[6]
13. [S Levine's lawyer [VP announced [NP the plea bargain] [PP in a press conference]

14. [S[S The judge adjourned the hearing] [PP in the afternoon]]

Each of these sentences takes the canonical form Subject-Verb-Object-PP. The task for the processing system is to determine whether the PP modifies the object (i.e., the PP is a constituent of the NP, as in (12)), the verb (i.e., the PP is a constituent of the VP, as in (13)), or the sentence (i.e., the PP is an adjunct to S, as in (14)). Some deny the need for a distinction between VP and S-modification. The difference is that with S-modification, the predication expressed by the PP has scope over the subject, while in VP-attachment it does not. For example, in (15), "in the park" applies to Levine, while in (16), "with 1,000 dollars" does not apply to Levine.

15. Levine made the announcement in the park.
16. Levine bought the stock with 1,000 dollars.

A number of solutions for the problem presented by post-verbal prepositional phrases have been offered. The most common techniques depend on structural (Frazier and Fodor 1978), semantic (Ford, Bresnan, and Kaplan 1982), or pragmatic (Crain and Steedman 1985) tests. MODL (McCord 1987) employs a combination of syntactic and semantic information for PP attachment. Independently, we formulated a preference strategy for PP attachment which uses ontological, generic and syntactic information to cover 99% of the cases in an initial test corpus, and which is 93% reliable across a number of types of text. This is the preference strategy we employ in KT. The PP attachment rules make use of information about the verb, the object of the verb, and the object of the preposition. A set of global rules is applied first, and if these fail to find the correct attachment for the PP, a set of rules specific to the preposition are tried. Each of these latter rules has a default. The global rules are stated informally in (17). with example sentences.

17a. time(POBJ)→ s_attach(PP)
If the object of the preposition is an expression of time, then S-attach the PP.
The judge adjourned the hearing in the afternoon.

b. lexical(V+Prep)→ vp_attach(PP)
If the verb and preposition form a lexicalized complex verb, then VP-attach the PP.
The defense depended on expert witnesses.

c. Prep=of→ np_attach(PP)
If the preposition is of *then NP-attach the PP.*
The ambulance carried the victim of the shooting.

d. intransitive(V) & motion(V) & place(POBJ) → vp_attach(PP)
If the verb is an intransitive verb of motion and the object of the preposition is a place then VP-attach the PP.
The press scurried about the courtroom.

e. intransitive(V) & (place(POBJ) OR temporal (POBJ) OR abstract(POBJ)) → s_attach(PP)
If the verb is intransitive and the object of the preposition is a place, temporal, or abstract, then S-attach the PP.
Levine worked in a brokerage house.

f. epistemic(POBJ)→ s_attach(PP)
If the prepositional phrase expresses a propositional attitude, then attach the PP to the S.
Levine was guilty in my opinion.

g. xp(...Adj-PP...)→ xp_attach(PP)
If PP follows an adjective, then attach the PP to the phrase which dominates and contains the adjective phrase.
Levine is young for a millionaire.

h. measure(DO)→ np_attach(PP)
If the direct object is an expression of measure, then NP-attach the PP.
The defendant had consumed several ounces of whiskey.

i. comparative→ np_attach(PP)
If there is a comparative construction, then NP-attach the PP.
The judge meted out a shorter sentence than usual.

j. mental(V) & medium(POBJ)→ vp_attach(PP)
If the verb is a verb of saying, and the object of the preposition is a medium of expression then VP-attach the PP.
Levine's lawyer announced the plea bargain on television.

Example 14 is handled by global rule (17a). Example 13 is handled by global rule (17j). The global rules are inapplicable with example 12, so the rules specific to "about" are called. These are shown below.

18a. intransitive(V) & mental(V)→ vp_attach(PP)
If the verb is an intransitive mental verb, then VP-attach the PP.
Levine spoke about his feelings.

b. Elsewhere→ np_attach(PP)
Otherwise, NP-attach the PP.
The government had uncovered an entire file about the scheme.

As a further example, the specific rules for "by" uses both generic (19a,b) and ontological (19c) knowledge.

19 a. nom(DO)→ np_attach(PP)
If the direct object is marked as a nominalization, then NP-attach the PP.
The soldiers withstood the attack by the enemy.

b. location(DO,POBJ)→ np_attach(PP)
If the relation between the direct object and the object of the preposition is one of location, then NP-attach the PP.
The clerk adjusted the microphone by the witness stand.

c. propositional(DO) & sentient(POBJ)→ np_attach(PP)
If the direct object is propositional and the object of the preposition is sentient, then NP-attach the PP.
The judge read out the statement by Levine.

d. Elsewhere→ s_attach(PP)
Otherwise, S-attach the PP.
The lawyers discussed the case by the parking lot.

These PP-attachment preference rules are remarkably successful when applied to real examples from actual text. However, they are not foolproof and it is possible to construct counterexamples. Take the global rule illustrated in (17a). We can construct a counterexample as in (20).

20. John described the meeting on Jan. 20th.

Sentence 20 is ambiguous. Jan. 20th can be the time of the describing or the time of the meeting. Perhaps there is a slight intuitive bias toward the latter interpretation, but the rules will assign the former interpretation. This is a counterexample only because "meeting" is a TEMPORAL noun and can plausibly have a time feature. Compare (21), which is identical except for the ontological attachment of the direct object and which is handled correctly by the global rule.

21. John described the proposal on Jan. 20th.

The problem of the interpretation of event nominals is a research topic we are working on.

The PP attachment rules are applied in the module DISAMBIG, which produces a disambiguated parse from the output of the MODL parser. The first step is to identify the sentence elements that form their inputs for each clause using **find_args.** The output of **find_args** is a list of the the direct object, object of the preposition, preposition, and main verb of the clause and the index of the clause (main, subordinate, and so on). The PP-attachment rules are in place and functional for one post-verbal prepositional phrase. Where more than one post-verbal prepositional phrase occurs, the current default is to attach the second PP to the NP of the first PP. However, this will not get the correct attachment in cases like the following.

22. The judge passed sentence on the defendant in a terse announcement.

A planned extension of the PP-attachment functionality will attack this problem by also keeping a stack of prepositions. The top of the stack will be the head of the rightmost PP. The attachment rules will be applied to PPs and the other constituents in a pairwise fashion until all are attached.

3.2 WORD SENSE DISAMBIGUATION

The word sense disambiguation method used in the system is a **combined local** ambiguity reduction method (Dahlgren 1988b). The method is local because word senses are disambiguated cyclically, from the lowest S-node up to the matrix node. Only when intrasentential sources of information fail are other sentences in the text considered by the disambiguation method. The algorithm is combined because it employs three sources of information. First it tries fixed and frequent phrases, then word-specific syntactic tests, and finally naive semantic relationships in the clause. If the fixed and frequent phrases fail, the syntactic and naive semantic rules progressively reduce the number of senses relevant in the clausal context. The algorithm was developed by considering concordances of seven nouns with a total of 2,193 tokens of the nouns, and concordances of four verbs with a total of 1,789 tokens of the verbs. The algorithm is 96% accurate for the nouns in these concordances, and 99% accurate for the verbs in these concordances.

Fixed phrases are lists of phrases that decisively disambiguate the word senses in them. For example, the noun "hand" has 16 senses. Phrases such as "by hand," "on hand," "on the one hand" have only one sense.

Syntactic tests either reduce the number of relevant senses, or fully disambiguate. For nouns, syntactic tests look for presence or absence of the determiner, the type of determiner, certain prepositional phrase modifiers, quantifiers and number, and noun complements. For example, only five of the 16 senses of "hand" are possible in bare plural noun phrases. For verbs, syntactic tests include the presence of a reflexive object, elements of the specifier, such as particular adverbs the presence of a complement of the verb and particular prepositions. For example, the verb "charge" has only its reading meaning "indict" when there is a VP-attached PP where the preposition is "with" (as determined by the prepositional phrase attachment rules). Syntactic tests are encoded for each sense of each word. The remainder of this section will illustrate disambiguation using naive semantic information and give examples of the naive semantic rules. (The complete algorithm may be found in Dahlgren 1988a).

3.2.1 NOUN DISAMBIGUATION

Naive semantic information was required for at least a portion of the disambiguation in 49% of the cases of nouns in the concordance test. Naive semantic inference involves either ontological similarity or generic relationships. Ontological similarity means that two nouns are relatively close to each other in the ontology, both upwards and across the ontology. If there is no ontological similarity, generic information is inspected. Generic information for the ambiguous noun, other nouns in the clause, the main verb, often disambiguate an ambiguous noun.

Ontological similarity is tested for in several syntactic constructions: conjunction, nominal compounds, possessives, and prepositional phrase modifiers. Many of the 16 senses of "hand" are ontologically distinct, as shown in Table 7.

1.	HUMAN	human body part
2.	DIRECTION	right or left
3.	INSTRUMENT	by hand
4.	SOCIAL	power, authority
5.	TEMPORAL	applause
6.	ROLE	laborer
7.	ARTIFACT	part of a clock

Table 7. Some senses of *hand*

In (23), only the HUMAN and ROLE senses (1 and 6) are possible, by ontological similarity. Generic knowledge of the verb "clear" is inspected for the final disambiguation to sense 6.

23. The farmer and his hand cleared the field.

In contrast, in (24), the relevant senses of "hand" are the HUMAN and ARTIFACT senses (1 and 7).

24. His arm and hand were broken.

At the point in the algorithm where naive semantic tests are invoked, syntactic tests have already eliminated the ARTIFACT sense, which does not occur with a personal pronoun. Thus the HUMAN sense is selected. In (25), only the ARTIFACT sense (7) is possible, by ontological similarity of "clock" and "hand."

25. The clock hand was black.

In (26), again the HUMAN and ROLE senses (1 and 6) are the only relevant ones by ontological similarity. Selection restrictions on "shake" complete the disambiguation.

26. John shook the man's hand.

In (27), sense 4 is selected because "affair" and sense 4 are both SOCIAL.

27. John saw his hand in the affair.

In (28), sense 1 is selected because both sense 1 and a sense of "paper" are attached to PHYSICAL.

28. The judge had the paper in his hand.

The word sense disambiguation algorithm tests for generic relationships between the ambiguous noun and prepositional phrases modifiers, adjective modifiers, and the main verb of the sentence. Two of the nine senses of "court" are shown in Table 8. In "the court listened to testimony," generic information for the second sense of "court" can be used to select sense 2. The generic information includes knowledge that one of the functions of courts has to do with testimony. In (29), sense 1 of "court" is selected because the generic representation of "court" contains information that witness stands are typical parts of courtrooms.

court1	PLACE	Typically, it has a bench, jury box, and witness stand. Inherently its function is for a judge to conduct trials in. It is part of a courthouse.
court2	INSTITUTION	Typically, its function is justice. Examples are the Supreme Court and the superior court. Its location is a courtroom. Inherently it is headed by a judge, has bailiffs, attorneys, court reporters as officers. Participants are defendants, witnesses and jurors. The function of a court is to hear testimony, examine evidence and reach a verdict. It is part of the justice system.

Table 8. Generic Information for Two Senses of *court*[7]

29. The witness stand in Jones's court is made of oak.

In (30), the adjective "wise" narrows the relevant senses from nine to the two INSTITUTION senses of "court."

30. The wise court found him guilty.

Generic knowledge of one sense of the verb "find" is then used to select between the court-of-law sense (2) and the-royal-court sense (4). Verb selection restrictions are powerful disambiguators of nouns, as many computational linguists have observed. In (31), the verb

charge1	Typically, if someone is charged, next they are indicted in court, convicted or acquitted. They are charged because they have committed a crime or the person who charges them suspects they have. Inherently, the charger and chargee are sentient, and the thing charged with is a crime.
charge2	Inherently, if someone charges that something is true, that someone is sentient, his goal is that it be known, and the something is bad.
charge3	Typically, if someone charges someone else an amount for something, the chargee has to pay the amount to the charger, and the chargee is providing goods or services. Inherently, the charger and chargee are sentients, the amount is a quantity of money, and the goal of the charger is to have the chargee pay the amount.

Table 9. Generic Information for Three Senses of *charge*

"last" requires a TEMPORAL subject, thus disambiguating "hand."

31. They gave him a hand which lasted 10 minutes.

3.2.2 VERB DISAMBIGUATION

Just as selectional restrictions of verbs disambiguate nouns, their arguments are powerful disambiguators of verbs. Subject, object, and oblique arguments are all taken into account by the algorithm. (32) and (33) illustrate the way that objects can disambiguate the verb "charge," generic entries for which are shown in Table 9.

In (32) the SENTIENT object selects sense 1. In (33), the MONEY object selects sense 3.

32. The state charged the man with a felony.
33. The state charged ten dollars for the fine.

The verb "present" can be disambiguated by its subject. It has at least three senses:

34. present1—"give"
present2—"introduce"
present3—"arrive in the mind of"

Senses 1 and 2 require SENTIENT subjects, so the third sense is selected in (35).

35. The decision presented a problem to banks.

(36) illustrates subject and object disambiguation. The SENTIENT subject narrows the possibilities to senses 1 and 2, and the "give" sense (1) is selected because it requires a PHYSICAL object argument.

36. John presented a bouquet to Mary.

3.2.3 DISAMBIGUATION RULES

The disambiguation method first tries fixed and frequent phrases, then syntactic tests, and finally naive semantic information. Each set of rules reduces the number of relevant senses of a word in the sentential (and extrasentential) context. There is a fixed set of commonsense rules for nouns and another one for verbs. They are tried in an order that inspects the main verb of the clause last, because the main verb often chooses between the last two relevant senses. An example of a noun rule is **ppmod,** which considers an ambiguous noun in relation to the head of a prepositional phrase modifier attached to the same higher NP as the ambiguous noun. There are two versions of the rule, one which looks for ontological similarity between senses of the ambiguous noun and the head of the PP, and one which looks for generic relationships between them. The output of **find_args** (in DISAMBIG, see Section 3.1) is used as a simplified syntactic structure inspected by these rules. This provides information as to whether or not a head noun is modified by a prepositional phrase. In the first version of the rule, the ontological attachment of the head of the PP is looked up and then senses of the ambiguous word with that same ontological attachment are selected. S1 is the list of senses of the ambiguous noun still relevant when the rule is invoked. S2 is the list of senses reduced by the rule if it succeeds. If the first version fails, the second is invoked. It looks for a generic relationship between senses of the ambiguous word and the head of the PP.

```
37. ppmod(Ambig_Word,{..Ambig_Word,Prep,
        Noun...},S1,S2) ←
        ontattach(Noun, Node) &
        ontselect(Ambig_Word,Node, S1,S2).
    ppmod(Ambig_Word,{..Ambig_Word,Prep,
        Noun...},S1,S2) ←
        generic_relation(Noun,Ambig_Word).
```

3.3 QUANTIFIER SCOPING

The semantic module of the KT system is capable of generating alternative readings in cases of quantifier scope ambiguities, as in the classic case illustrated in (38).

38. Every man loves a woman.

In this example either the universally quantified NP ("every man") or the existentially quantified NP ("a woman") can take widest scope. In such sentences, it is generally assumed that the natural scope reading (leftmost quantified expression taking widest scope) is to be preferred and the alternative reading chosen only under explicit instructions of some sort (such as input from a user, for example, or upon failure to find a proper antecedent for an anaphoric expression). Under this assumption, in a sentence like (39), the indefinite NP would preferentially take widest scope.

39. A woman loves every man.

But there are a number of cases similar to (39) where an expression quantified by "every" appears to the right of an indefinite NP and still seems to take widest scope. Ioup (1975) has discussed this phenomena, suggesting that expressions such as "every" take widest scope inherently. Another computational approach to this problem is to assign precedence numbers to quantifiers, as described in McCord (1987). However, our investigation has shown that commonsense knowledge plays at least as large a role as any inherent scope properties of universal quantifiers.

Consider (40). In the natural scope reading, "every lawyer" takes scope over "a letter" and we have several letters, one from each of the lawyers, i.e., several tokens of a one-to-one relationship. In the alternative reading, "a letter" takes scope over "every lawyer" and we have only one letter, i.e., one token of a many-to-one relationship relationship. Both scope readings are plausible for (40).

40. The judge read a letter from every lawyer.

In (41) only the alternative reading (several tokens of one-to-one) is plausible.

41. The politician promised a chicken in every pot.

In (42), however, only the natural reading (one token of many-to-one) is plausible.

42. The prince sought a wife with every charm.

Even in (40), however, speakers prefer the one-to-one relationship, the alternative reading. That is, speakers prefer the reading that denotes several tokens of a one-to-one relationship (several letters) over one which denotes one token of a many-to-one relationship unless there is strong commonsense knowledge to override this preference. We know that in our culture princes can have only one wife, so in the case of (42) speakers prefer one token (one wife) of many-to-one to several tokens of one-to-one. Similar arguments apply to the following examples (43)–(45), which correspond to (40)–(42) respectively.

43. A judge decided on every petition.
44. A lawyer arrived from every firm.
45. A company agent negotiated with every union.

Thus, if there is an inherent tendency for universal quantifiers to take widest scope where scope ambiguity is possible, it derives from the human preference for viewing situations involving universally quantified entities as many tokens of one-to-one relationships.[8] In KT, first we prefer wide scope for the universal quantifier. In cases where this is not the natural scope interpretation, (i.e., the universal quantifier is to the right of a containing existentially quantified NP), we can use facts encoded in the generic data base to override the preference. For example, when processing (42) we would discover that a man may have only one wife. The generic entry for "wife" tells us that "wife" is a role in a "marriage." The generic entry for "marriage" tells us that "marriage" is a relation between exactly one husband and exactly one wife. This knowledge forces the many-to-one interpretation. The cases where this is necessary turn out to be rare. Curiously, the preposition "with" correlates very highly with many-to-one relationships. Thus our strategy for the present has been to consider overriding the preference only when the universal quantifier is in an NP which is the object of "with." In these cases we access the generic knowledge, as described above.

3.4 OPAQUE CONTEXTS

In KT, clauses in opaque contexts (embedded under propositional attitude verbs such as "hope," "believe," "deny") are handled by asserting the predicates generated from the clause into partitioned databases, which correspond to the delineated DRSs of Asher (1987). Each partition is associated with the speaker or the individual responsible for the embedded clause. Reasoning can then proceed taking into account the reliability and bias of the originator of the partitioned statements, as in Section 2.2.4.1 An example follows.

46. *Text:* Meese believes that Levine is guilty.
Textual Database: believe(s1,meese,p1)
Partition: p1 :: guilty(s2,levine)

3.5 MODALS[9]

The English modals *can*, *may*, *must*, *will*, and *should* are high-frequency items in all kinds of texts. They can be easily parsed by a single rule similar to the rules that handle auxiliary "have" and "be" because all the modals occupy the same surface syntactic position (i.e., the first element in the auxiliary sequence). However, the modals present some considerable problems for semantic interpretation because they introduce ambiguities and induce intensional contexts in which possibility, necessity, belief, and value systems play a role. In the KT system, we are concerned with what is known by the system as a result of reading in a modal sentence. In particular we are interested in what status the system assigns to the propositional content of such a sentence.

To illustrate the problem, if the system reads "Levine engaged in insider trading," then an assertion can justifiably be added to the knowledge base reflecting the fact that Levine engaged in insider trading. The same is true if the system reads "The Justice Department knows that Levine engaged in insider trading." But this is not the case if the system reads "The Justice Department believes that Levine engaged in insider trading." In this case the statement that Levine engaged in insider trading must be assigned some status other than fact. Specifically, since "believe" introduces an opaque context, the propositional content of the embedded clause would be assigned to a partitioned data base linked to the speaker *the Justice Department*, as described in the previous section. A similar problem exists in modal sentences such as "Levine may have engaged in insider trading."

There are two types of modal sentences. In Type I modal sentences the truth value of the propositional content is evaluated with respect to the actual world or a set of possible other states of the actual world directly inferable from the actual world. Examples are

47. Levine must have engaged in insider trading.
48. The Justice Department will prosecute Levine.
49. Levine can plead innocent.

In Type II modal sentences, we say that a second speech act is "semantically embedded" in the modal sentence. The modal sentence is successful as an assertion just in case the secondary speech act is in effect in the actual world. In these cases the truth value of the propositional content is evaluated with respect to some set of deontic or normative worlds. The modal is viewed as a quantifier *cum* selection function. Thus, for a sentence of the form $\mu = NP\ Modal\ VP$, μ is true in the actual world just in case *NP VP* is true in at least one/every world (depending on Modal) in the set of deontic or normative worlds selected by Modal. Examples are

50. Levine must confess his guilt.
51. Levine may make one phone call.
52. Levine should get a good attorney.

In (50) a command is semantically embedded in the assertion; in (51) a permission is semantically embedded in the assertion; and in (51) the issuance of a norm is semantically embedded in the assertion.

Type I modal sentences are of assertive type according to the speech act classification scheme in Searle and Vanderveken (1985). These include the standard assertions, reports, and predictions, and a proposed new type, **quasi-assertion**. They must be distinguished in a text-understanding system from Type II modal sentences that embed other types of speech acts, because only Type I modal sentences make a contribution to the textual knowledge base. In addition, some modal sentences are ambiguous between Type I and Type II, for example (47), (51), and (50).

For the KT system, disambiguating the ambiguous modals "may" and "must" results in changing the syntactic input to the semantic module. The surface syntactic parse that is output by the parser is converted into the equivalent logical form where the epistemic (Type I) uses of ambiguous modals are treated as sentential operators and the nonepistemic uses of ambiguous modals (Type II) are treated as modifiers of the verb. Sentences containing ambiguous modals can be assigned the correct status by a simple disambiguation algorithm that makes appeal to the ontological classification of the main verb, specifically whether or not the verb is STATIVE, following Steedman (1977). Disambiguation takes place in DISAMBIG, the same module that converts the labelled bracketing to a modified parse for input to the DRT semantic translator. At this point, a determination is made whether the modal is a sentential operator or a modifier of the verb. The propositional content of quasi-assertions and predictions can be added directly to the dynamically constructed textual data base if they are appropriately marked with probability ratings. On this view, "will" and one sense of "may" are taken as denoting strong and weak prediction and are not viewed as tenses. That is, when using these modals, the speaker is indicating his confidence that there is a high/moderate probability of the propositional content being true in the future. In the present state of the system, every predicate contains a tense argument and there is a tense predicate relating every tense argument to a tense value (such as "pres.," "future," etc.).[10] In a planned extension to the system these tense predicates will also contain probability ratings. For example, given the continuum of speaker commitment to the truth of the statement "Levine engaged in insider trading" illustrated in (53), we would have the corresponding predicates in the DRS shown in (54).

53. Full Assertion: *Levine engaged in insider trading*
    Strong Quasi-Assertion: *Levine must have engaged in insider trading*
    Weak Quasi-Assertion: *Levine may have engaged in insider trading*
54. Full Assertion: engage(e1,levine), tense(e1,past,1)
    Strong Quasi-Assertion: engage(e1,levine), tense(e1,past,0.9)
    Weak Quasi-Assertion: engage(e1,levine), tense(e1,past,0.5)

This hierarchy reflects the "epistemic paradox" of Karttunen (1971), in which he points out that in standard modal logic *must(P)* or *necessarily, P* is stronger than plain assertion whereas *epistemically-must(P)* is weaker than plain assertion. This results from the fact that the standard logic necessity operator quantifies over every logically possible world, plain assertion of *P* is evaluated with respect to the actual world, but the epistemic modal operator quantifies only over the epistemically accessible worlds, a set which could possibly be null.

Assertions of possibility (49) trigger the inferring of enabling conditions. For any event there is a set of enabling conditions that must be met before the event is possible. For John to play the piano, the following conditions must be met:

55. 1. John knows how to play the piano.
    2. John has the requisite permissions (if any).
    3. A piano is physically available to John.
    4. John is well enough to play.
    5. . . .

These can be ordered according to saliency as above. This, we claim, is why the sentence "John can play the piano" most often receives the interpretation (1), less often (2), and practically never (3) or (4) unless explicitly stated, as in "John can play the piano now that his mother has bought a new Steinway." The enabling conditions are encoded in KT as part of the generic representation for verbs. When a modal sentence is interpreted as a full assertion of possibility (*poss(p)*), this triggers the inference that the most salient of the enabling conditions is in fact true. The difference between *poss(p)* and *p* being processed for KT, is that if *p* is output, then *p* is added to the textual database and the most salient enabling condition is also inferred. But if *poss(p)* is output, then only the most salient enabling condition is inferred, but *p* is not added to the textual database. Notice that this simply reflects the fact that if I say "John can play the piano," I am not saying that John is playing the piano at that very moment.

Type II modal sentences present a more complex problem for interpretation. The commands, permissions, and norms reported in Type II modal sentences are asserted into partitioned databases in the same way as clauses in opaque contexts. The only difference is that in most cases the issuer of the command, permission, or norm reported in a modal sentence is not known. Semantic translation in DRT proceeds via categorial combination. By the time the modal sentence reaches the DRT module, the semantic type of the modal is unambiguous and the appropriate lexical entry can be retrieved. The creation of appropriate predicates to express the variety of modal statements is the task of the DRT module.

4 The Query System

4.1 Open and Closed World Assumptions

It is well known that negation-as-failure is a sound extension of SLD resolution only when the database is complete, i.e., when it represents a closed world (Lloyd 1984). Since our databases will always include some predicates about which we have only incomplete information, we cannot assume a completely closed world. The open world assumption, though, makes it unsound to use the very useful negation-as-failure rule. Fortunately, it is well known that we can restrict the use of this rule to just those predicates known to be complete, keeping an open world assumption for other predicates. We accordingly specify that some of our general knowledge comprises a complete, closed world in the appropriate sense, but we do not make this assumption about textual knowledge.

4.2 FUNCTIONING OF THE QUERY SYSTEM

Queries are handled just like text up through conversion to FOL. The FOL form of the query is then passed to REASONER, which decides which database is the most likely source of the answer. REASONER can access the textual database, the verb and noun generic databases, and the ontology. The reasoning is complex and dependent on whether the query form is ontological, factual, or generic. A search sequence is then initiated depending on these factors. The form of the answer depends on the search sequence and the place where the answer is found.

4.2.1 ANSWERS

The form of the answer depends on the reliability of the knowledge encoded in the database where the answer is found. The text is considered authoritative. If an answer is found in the text, search is terminated. The ontology is considered a closed world (see discussion above). This means that yes/no ontological questions are answered either "yes" or "no." The textual and generic databases are considered an open world. If an answer is not found, and no further search is possible, the system concludes, "I don't know." Answers found in the generic databases are prefaced by "Typically," for information found in the first list (of typical features) or "Inherently," for answers found in the second list (of inherent features).

4.2.2 QUESTION TYPES

An ontological question is in a copular sentence with a non-terminal ontological node in the predicate.

56. Ontological Questions:
    Is a man human?
    Is the man a plant?

A factual question is one couched in the past tense, present progressive, and/or where the subject is specific (a name or a definite NP). Specific NPs are assumed to have already been introduced into the system. Our simplified definition of generic question is one which contains an inherently stative verb or a non-stative verb in the simple present combined with a non-specific subject (indefinite NP).[11]

57. Factual Questions:
    Did John buy a book?
    Is the man happy?
    Who bought the book?

    Generic Questions:
    Does a man buy a book?
    Does the man love horses?

Ontological questions are answered by looking in the ontological database only. If an answer is not found, the response will be "no" if the query contained an ontological predicate (such as PLANT or ANIMAL) because the ontology is a closed world.

58. Text: John is a man who bought a book.
    Is a plant living?—Yes.
    Is the man an animal?—Yes.
    Is the man a plant?—No.

Factual queries (non-generic questions) go to the textual database first. If an answer is not found, then the generic knowledge is consulted. If an answer is found there, the appropriate response (Typically.., Inherently..) is returned. Otherwise the response is, "I don't know."

59. Text: John is a man who bought a book.
    Did John buy a book?—Yes.
    Who bought a book?—John.
    Is John the President?—I don't know.
    Does John wear pants?—Typically so.
    Where did John buy the book?—Typically, in a store.

Generic queries go only to the generic database.

60. Does a man wear pants?—Typically so.
    What is the function of a lawyer?—Inherently, represents clients.
    Is an apple red?—Typically so.
    Does a man love flowers?—I don't know.
    Where does a man buy a book?—Typically, in a store.
    Who buys a book?—Inherently, a sentient.

In addition to these general rules, there is special handling for certain types of queries. Questions of the form "Who is . . ." and "What is . . ." are answered by finding every predicate in any data base that is true of the questioned entity. For example, in any of the examples above, the question "Who is John?" would trigger a response that includes a list of all the nodes in the ontology which dominate the position where "John" is attached plus the information that John is a man, is tall, and bought a book. The question "What is a vehicle?" would trigger only the list of ontological nodes because there is no specific vehicle in the domain of this example. Questions such as "Who buys a book?" and "What is driven?" are answered by stating selectional restrictions on the verb—*Inherently, a sentient buys a book,* and *Inherently, a vehicle is driven.*

Finally, it is possible to override the generic information in specific cases while still retaining the capability of accessing the generic information later, as the following example shows.

61. What color is an airplane?—Typically, white.
John bought an airplane.—OK.
What color is the airplane. - The text does not say, but typically white.
The airplane is red.—OK.
What color is the airplane?—The text says red.
What color is an airplane?—Typically, white.

Thus the system uses default logic (Reiter 1980).

REASONER, therefore, is sensitive to a number of factors which make the system seem to understand the queries in a natural way. The responses generated also reflect the continuum of reliability of information which is available to a human reasoner. A flow chart of the search strategies in REASONER is shown in Figure 3.

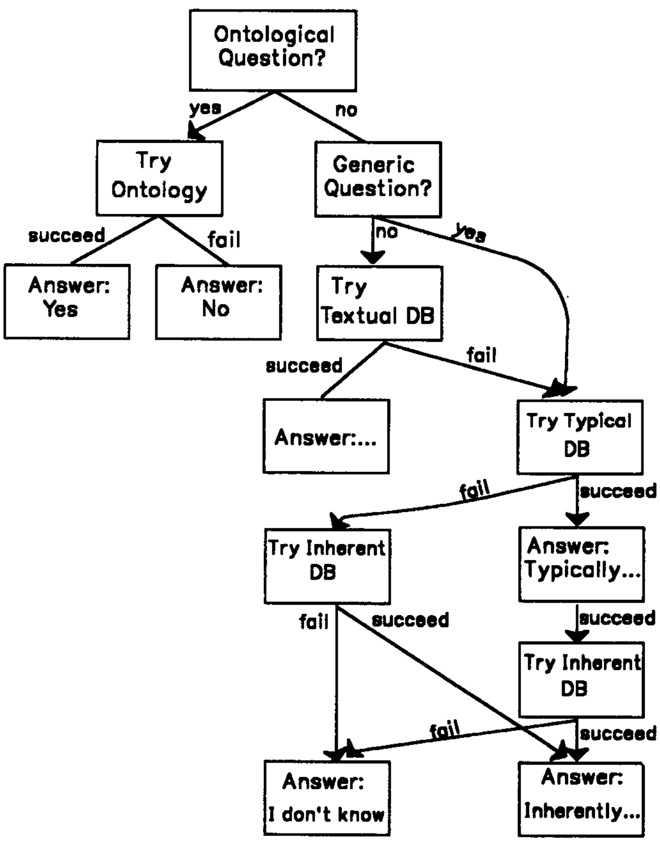

Figure 3. The Query System

5 Naive Semantics and Discourse Phenomena[12]

Most computational treatments of discourse phenomena acknowledge the role of world knowledge in anaphora resolution, temporal reasoning, and causal reasoning (Reichman 1985; Grosz and Sidner 1986; Wada and Asher 1986). However, in the past the only method for encoding and incorporating world knowledge involved writing a detailed script for every real-life situation, directly encoding the probable sequence of events, participants, and so forth (Schank and Abelson 1977). This section will demonstrate that word level naive semantics offers a principled, transportable alternative to scripts. NS is a powerful source of information in discourse reasoning. Along with syntactic, compositional semantic, and discourse cue information, NS can be used to reason heuristically about discourse and drive many of the inferences drawn by people when they read a discourse. The role of syntax and compositional semantics will be underplayed in what follows, only because these contributions have been thoroughly treated by others (Reinhart 1982; Asher and Wada 1988; Kamp 1981; Grosz and Sidner 1986; Reichman 1985; Webber 1985).

5.1 Anaphora

In anaphora resolution, syntactic constraints, accessibility (in the sense of Kamp 1981), and discourse segmentation work in concert to limit the number of antecedents available to an anaphoric pronoun or definite NP. However, it is clear that the resultant saliency stack can end up with more than one member (Asher and Wada 1988). It is at this point in the reasoning that naive inference is required. Consider the following.

62. Levine's friend is a lawyer. *He* won *his* case for *him*.

Syntactic rules exclude a reading in which "he" and "him" co-refer. Since both Levine and the lawyer are possible antecedents of all of the pronouns and therefore are both in the saliency stack, the following readings remain:

63. i. Levine won Levine's case for the lawyer.
ii. Levine won the lawyer's case for the lawyer.
iii. The lawyer won Levine's case for Levine.
iv. The lawyer won the lawyer's case for Levine.

Reading (iii) is the one people select. Generic knowledge of "lawyer" suffices to make this selection. The generic representation of "lawyer" in (9) includes information that lawyers argue cases. An inspection of generic knowledge for "argue" reveals that one goal of arguing is winning. Using these two facts, the system can infer that the lawyer won, so the lawyer is the subject of the sentence. Thus the benefactee, by disjointness, must be Levine. Now the feature of "lawyer" that says that lawyers represent their clients could be used to figure that the case is Levine's.

Turning to definite anaphora, a definite description often has an implicit antecedent in the discourse, rather than an explicit or deictic one. Unless world knowledge is used, it is impossible to recover these. Consider the following.

64. Levine's trial was short. The testimony took only one day.

A generic representation of "trial" is shown below. Using it, the antecedent of "testimony" in (65) can be identified as a part of the trial.

65. trial -

Typically, in a trial first there are oaths, statements and testimony, then the judge instructs the jury, the jury deliberates and announces a verdict, and the judge pronounces the sentence. There are roles of witnesses and spectators. A trial lasts about three days.

Inherently, a trial has roles of judge, clerk, bailiff, attorneys, plaintiff, and defendant. The goal of a trial is to settle a dispute or determine the guilt or innocence of a defendant.

An interesting subset of definite anaphora is definite event anaphora (Asher 1988). In (66), we know that "decision" refers to an event, because it is a deverbal nominal. Generic knowledge for deverbal nominals is the same as for verbs. So we know that there is some definite SOCIAL MENTAL event. Both commenting and sentencing are SOCIAL MENTAL verbs. Generic knowledge of the verb "sentence" includes knowledge that a sentencing is enabled by a decision. This can be used to infer that the antecedent of "the decision" is the sentencing event reported in the first sentence, rather than the commenting event. Thus $e_4 = e_1$.

66. (e_1) The judge sentenced Levine to a short term.
(e_2) He commented that Levine had been cooperative.
(e_3) The decision (e_4) surprised an attorney.

If the generic knowledge has an implication which is similar to the event nominal, the correct antecedent can be inferred. The resulting DRS is shown in (67).

67.

x_1, x2, a_1, e_1, x_3, p, e_4, x_3, a_2, e_2, e_3
----
judge(x_1)
levine(x_2)
term(a_1)
short(a_1)
e_1 sentence(x_1,x_2)
to(e_1,a_1)
he(x_3)
e_2 comment(x_3,p)
p: s_2
----
s_2 cooperative(x_2)

$x_3 = x_1$
decision(e_3)
attorney(a_2)
e_2 surprise(e_3,a_2)
$e_3 = e_1$

5.2 TEMPORAL REASONING

Temporal reasoning involves the assignment of relationships among the times of the events reported in the discourse. Following the Reichenbachian approach to tense found in Partee (1984), there are three elements to temporal reasoning: event time (e_i), reference time (r_i), and speech time (now). In a sequence of simple past tense clauses, the reference time is updated each time. So in text (66), the temporal equations look as follows.

68. $r_1 <$ now.
$e_1 \subseteq r_1$
$r_1 < r_2$
$e_2 \subseteq r_2$
$r_2 < r_3$
$e_3 \subseteq r_3$

However, the reference time is not always updated by a simple past tense verb. In addition to the effects of clause aspect, which will be discussed below, commonsense knowledge affects temporal reasoning. Two events reported in a sequence of simple past tense clauses can overlap in time, or the second event (in textual sequence) can occur before the first. Naive semantic representations are sufficient to assign these relations correctly in many cases. The typical implications of events in verb representations can be used to infer overlap. Consider the following discourse.

69. (e_1) Levine made a statement. (e_2) He said he was sorry.

We know that a "statement" is a deverbal nominal of "state," and that the goal of "state" and of also "say" is communicating ideas. This knowledge is used to infer that $e_2 \subseteq e_1$.

The regular temporal presuppositions of implicational features on verbs are very powerful in inferring the relative order of events. An event e_i must occur before an event e_j in order to cause e_j.

cause(e_1,e_2)	e_2 before e_1
enable(e_1,e_2)	e_2 before e_1
goal(e_1,e_2)	e_2 after e_1
consequence(e_1,e_2)	e_2 after e_1
whnext(e_1,e_2)	e_2 after e_1
when(e_1,e_2)	e_2 overlap e_1
where(e_1,e_2)	e_2 overlap e_1
how(e_1,e_2)	e_2 overlap e_1

Table 10. Temporal Presuppositions of Verb Implications

Table 10 lists the temporal presuppositions of several generic features of verbs. This works well when one of the implications of a verb mentions another verb (or related verb) in the text, and tense or adverbial modifiers do not give clues as to the temporal relations. In (70), the fact that e_2 is a typical cause of e_1, can be used to infer that e_2 occurred before e_1.

70. (e_1) Levine was found guilty. (e_2) He broke the law.

Similarly, in (71), the fact that buying typically takes place in stores, can be used to infer that e_2 overlapped e_1 in time. (The stativity of the verb in the second sentence is also a weak, but insufficient indicator of the temporal relationship between the sentences).

71. (e_1) Levine bought a book. (e_2) He was in a store on 5th Ave.

Temporal properties of nouns also require naive semantics. Role terms have a temporal element (Enc 1987). A term such as "President of the U.S." has a DURATION feature in the naive semantic representation. This enables appropriate inferences concerning the occupants of roles. In (72), the system can infer that the individual who declared war on insider trading is probably not the same as the one whose office indicted Levine.

72. The Attorney General declared war on insider trading in 1964. The Attorney General's office indicted Levine six months ago.

Another way in which naive semantic representations can be used in temporal reasoning has to do with the assignment of tense to deverbal nominals. In (73), the nominal "violations" refers to events which can be placed in time before the sentencing. Naive semantic knowledge can be used to infer that the violations took place before the sentencing, because violations are illegal, sentencing is the final stage of a trial, a trial determines whether the defendant is guilty, and guilty parties have done something illegal.

73. Levine, who engaged in massive insider trading, was sentenced to two years in prison. His violations of the securities laws were shocking.

Another way that naive semantics helps temporal reasoning, comes from the assignment of coherence relations using naive semantics. How this can be done is discussed below (Section 5.4). Given certain rhetorical relations between two clauses, certain temporal relationships are indicated. Table 11 lists the temporal predications of certain coherence relations. The first group are relations which hold between events or states δ_1 and δ_2, where δ_1 occurs before δ_2. In (74) the system can recognize that e_1 is before e_2, once it assigns the relation Cause between them.

74. (e_1) Levine broke the law. (e_2) He was indicted.

The same thing would work if the sentences were in reverse order in the text. The second group in Table 11 lists the coherence relations that indicate that events or states overlap in time. The third group lists those relations that require that the source event or state is in the speech time (now).

δ_1 and δ_2—event or state reference markers
$r_1 < r_2$—reference times

1. $\delta_1 \subseteq r_1$, $\delta_2 \subseteq r_2$	Elaboration(δ_1,δ_2), Cause(δ_1,δ_2), Goal(δ_2,δ_1) Evidence(δ_2,δ_1), Enablement(δ_1,δ_2), Comment(δ_2,δ_1),
2. δ_1, $\delta_2 \subseteq r_1$	Parallel(δ_1,δ_2), Contrast(δ_1,δ_2), Description(δ_1,δ_2), Qualification(δ_1,δ_2), Evidence(δ_1,δ_2) Generalization(δ_1,δ_2), Import(δ_1,δ_2),
3. $\delta_2 \subseteq$ now	Import (δ_2,δ_1), Evaluation(δ_2,δ_1)

Table 11. Tense and Coherence Relations

5.3 CLAUSE ASPECT

Clause aspect refers to the aspect of an entire clause. The classical example in (75) through (77) illustrates differing clause aspects with the same verb and tense. (75) is telic, while (76) and (77) are activity clauses.

75. John pushed the cart under a shed.
76. John pushed the cart under adverse conditions.
77. John pushed the cart.

Clause-stative means not only the opposition between an inherently stative verb and an eventive verb (as in "be" vs. "hit"), but to the various ways in which a whole clause can end up being stative, as with the presence of the progressive, or a number of other factors. **Clause-telic** means that the clause reports a change of state with a terminus. "John built the house" is clause-telic, while "John was building the house" is clause-stative. A telic clause has an ACHIEVEMENT or ACCOMPLISHMENT verb not in the progressive, not in the simple present (which would be habitual, and with no modal (e.g., "John will build the house" is not telic). **Clause-activity** has to do with a clause which reports an event which has no terminus, and which has the sub-interval property (Bennett and Partee 1978), as "John ran." Naive semantic knowledge is used to assign clause aspect. In (75), naive semantic knowledge that a shed is a PLACE, and generic information of the relative sizes of carts and sheds, can be used to infer that the shed was a destination for the cart. An ACTIVITY verb such as "push," with a destination argument in the verb phrase results in a telic clause. This inference for (75) would not hold for (76). Similarly, an ACHIEVEMENT or ACCOMPLISHMENT verb indicates a TELIC clause, but other arguments can change them to ACTIVITY (cf. Moens and Steedman 1987). For example, in (78) the clause is TELIC, while in (79) it is ACTIVITY, and in (80) it is ambiguous between TELIC and ACTIVITY.

78. The prosecutor questioned the point.
79. The prosecutor questioned the witness for an hour.
80. The prosecutor questioned the witness.

5.4 COHERENCE RELATIONS

Coherence relations are handled in the KT system as added predicates in a cognitive DRS. In the DRS representing two clauses connected by a discourse cue word such as "because," a predicate cause(e_1,e_2) is represented. Similarly, where the first event typically causes the second, the system guesses the cause relation between the two event reference markers, and a cause predicate is introduced into the DRS, resulting in a cognitive (or inferred) DRS. Coherence relations are assigned using syntax, temporal relations, clause aspect, discourse cues, and naive semantics. An algorithm for coherence relation assignment has been developed (Dahlgren 1988c, 1989). This section will illustrate only the contribution of naive semantics and will not delve into the complex problem of the interactions among the several sources of information. Grosz and Sidner (1986) argue that coherence relations are not a useful analytical tool because no clear, closed set of them has been discovered. However, there is ample psycholinguistic evidence that in constructing the interpretation of a text, and in recalling what it said, coherence relations are inferred and used by readers (Rickheit and Strohner 1985). In terms of computational linguistics, coherence relations are useful for text summarization and relevance reasoning. In text summarization, only the general actions, not the elaborations, can be included in the summary. Descriptive and other background clauses can be ignored. Similarly, relevance can be inferred from the causal implications of events reported in a text. If a reader says that he or she wants to read about events that affect the construction industry, for example, and the typical consequence of some event in a text affects the construction industry, then that reader is interested in that text.

The naive semantic representations of nouns and verbs contain sufficient information to handle a large number of cases in which world knowledge is required to structure the discourse. Generic representations of the typical implications of verbs such as cause, goal, enablement, and consequence are the very same information as coherence relations. Their content means, "If there was an event (or state) of VERB1ing, then it probably had as goal a later event (or state) of VERB2ing." For example, "If there was an event of buying, it probably had as goal a state of owning." Naive semantic representations contain generalizations about objects and actions which are common to a linguistic community, and thus explain the ability to understand a discourse without resort to particular scripts describing familiar real-world situations.

Using generic and ontological representations derived from psycholinguistic data, coherence relations can be assigned. To infer goal(e_1,e_2) for text (81), knowledge that "profit" is money can be used to relate (e_2) to the the goal feature of "invest."

81. (e_1) John invested heavily. (e_2) He made a huge profit.

Some of the naive semantic representation of "invest" is shown below:

82. Investing is typically lucrative and is accomplished with money. Inherently, sentients do the investing with the goal of making money.

Similarly, in (83), the generic entry for "insider trading" can be used to infer that Levine broke the law. The entry associated with reading 1 of the verb "charge," as shown in Table 9, includes information that a typical cause of charging someone is that that someone has committed a crime. Putting these two together, cause(e_1,e_2) can be inferred.

83. (e_1) Levine engaged in insider trading.
    (e_2) The government charged him with violations of the securities laws.

The segmentation of discourse takes into account paragraphing, discourse cues such as *Turning to . . .*, *In summary*, clause aspect, temporal relations, and coherence relations. In this section we will briefly illustrate that naive semantics is one source of information in discourse segmentation. In narrative, a clear distinction can often be made between segments consisting of sequences of actions that are the foreground of the narrative, and segments that provide the background or setting for the action. If the author does not give clear discourse cues of the switch to a setting or situation segment, the shift can be inferred using naive semantics. In our method, discourse segments are related to each other the same way as clauses (as in Mann and Thompson [1987] and Hobbs [1985]), so the relationship here is one of situation_activity(Seg2,Seg1) where Seg1 is a sequence of actions. Consider (84).

84. Levine engaged in insider trading at his firm. He was charged and found guilty of violations of the securities laws. He was sentenced by Judge Goettel. *Levine was happy at his firm. The audience waited with baited breath to hear what Judge Goettel would say.*

In the text, there is a change of segment at "Levine was happy at his firm." The segment is a SITUATION-ACTIVITY segment. It describes what was going on when Levine was engaging in illegal practices. The change from a sequence of actions to a background segment is indicated by several factors, including the use of the stative and the place adverbial. Another factor is naive semantic knowledge of "firm." Working takes place at a firm, and this knowledge can be used to infer that "Levine was happy at his firm" refers to a long-term situation in which Levine was working. In that situation, he was happy. Thus the sentence is not about some specific action, but is a generalization about Levine's condition as a worker. Such a generalization indicates a change of segment from a sequence of actions to a SITUATION_ACTIVITY segment.

6 Conclusion

Naive semantics is a level of cognitive representation of concepts that can be discovered empirically and represented in a principled way with FOL without resort to a special knowledge representation language. Because NS representations are linked to ordinary words and do not depend on a special knowledge representation level, they should be transportable from one text to another. The KT system demonstrates that these rich representations are powerful in resolving many of the large residue of ambiguities that remain after the work of a purely syntactic parser is completed.

References

Abbott, V.; Black, B. J.; and Smith, E. 1985 The Representation of Scripts in Memory. *Journal of Memory and Language* 24(1): 179–199.

Ashcraft, M. H. 1976 Property Norms for Typical and Atypical Items from 17 Categories: A Description and Discussion. *Memory and Cognition* 6(3):227–232.

Asher, N. 1987 A Typology for Attitude Verbs and Their Anaphoric Properties. *Linguistics and Philosophy* 10:125–198.

Asher, N. 1988 The Semantics for Proposition-type Nominals and Proposition-type Anaphora. Paper presented at the University of Texas Conference on the Structure of Events and Natural Language Metaphysics.

Asher N. and Wada, H. 1988 A Computational Account of Syntactic, Semantic and Discourse Principles of Anaphora Resolution, Forthcoming in *Journal of Semantics*.

Bennett, M. and Partee, B. 1978 *Towards the Logic of Tense and Aspect in English*. Indiana University Linguistics Club: Bloomington, IN.

Boyd, R. 1986 Natural Kinds, Homeostasis and the Limits of Essentialism. Forthcoming.

Brachman, R. J.; Gilbert, V. P.; and Levesque, H. J. 1985 An Essential Hybrid Reasoning System. *Proceedings of the International Joint Committee on Artificial Intelligence. IJCAI* 1:532–539.

Chang, C.-L. and Lee, R.C.-T. 1973 *Symbolic Logic and Mechanical Theorem Proving*. Academic Press, New York, NY.

Crain, S. and Steedman, M. 1985 On Not Being Led Up the Garden Path: the Use of Context by the Psychological Syntax Processor. In *Natural Language Parsing*, David R. Dowty, Lauri Karttunen and Arnold M. Zwicky, (Eds.) Cambridge University Press, 320–358.

Dahlgren, K. 1985a The Cognitive Structure of Social Categories. *Cognitive Science* 9:379–398.

Dahlgren, K. 1985b Social Terms and Social Reality. *Folia Linguistica Historica* 6:107–126.

Dahlgren, K. 1988a *Naive Semantics for Natural Language Understanding*. Kluwer Academic Press, Boston, MA.

Dahlgren, K. 1988b Using Commonsense Knowledge to Disambiguate Word Senses. In: F. St. Dizier and V. Dahl, Eds., *Natural Language Understanding and Logic Programming 2*. North Holland, Amsterdam.

Dahlgren, K. 1988c Coherence Relation Assignment. *Proceedings of the West Coast Conference of Linguistics*, Fresno, CA.

Dahlgren, K. 1989 Coherence Relations and Naive Semantics. Paper given at the Symposium on Modelling Discourse Structure: Discourse Segments and Discourse Relations, University of Texas, Austin, TX.

Dahlgren, K. and McDowell, J. 1986a Kind Types in Knowledge Representation. *Proceedings of the 11th International Conference on Computational Linguistics* (COLING-86).

Dahlgren, K. and McDowell, J. 1986b Using Commonsense Knowledge to Disambiguate Prepositional Phrase Modifiers. *Proceedings of the American Association for Artificial Intelligence 86*.

Dowty, D. R. 1979 *Word Meaning and Montague Grammar*. D. Reidel Publishing Company, Dordrecht, Holland.

Enc, M. 1986 Towards a Referential Analysis of Temporal Expressions. *Linguistics and Philosophy* 9:405–426.

Fahlman, S. 1979 *NETL: A System for Representing and Using Real-World Knowledge*. MIT Press, Cambridge, MA.

Ford, M.; Bresnan, J.; and Kaplan, R. 1981 A Competence-Based Theory of Syntactic Closure. In: Joan Bresnan, Ed., *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge, MA.

Frazier, L. and Fodor, J.D. 1978 The Sausage Machine: A New Two-stage Parsing Model. *Cognition* 6:291–325.

Gabbay, D.M. and Sergot, M.J. 1986 Negation as inconsistency. *Journal of Logic Programming* 1:1–35.

Gelman, R. and Spelke, E. 1981 Thoughts about Animate and Inanimate Objects. In: J. H. Flavell and L. Ross (Eds.), *Social Cognitive Development*. Cambridge University Press, Cambridge, England.

Graesser, A. and Clark, L. 1985 *Structure and Procedures of Implicit Knowledge*. Ablex. Norwood, NJ.

Graesser, A. and Hopkinson, P. 1987 Differences in Interconcept Organization between Nouns and Verbs. *Journal of Memory and Language* 26.

Grosz, B. and Sidner, C. 1986. Attention, Intensions, and the Structure of Discourse. *Association for Computational Linguistics* 12:175–204.

Hobbs, J. 1985 On the Coherence and Structure of Discourse. CSLI Report # CSLI-85-37, CSLI, Stanford University, Stanford, CA 94305.

Huttenlocher, J. and Lui, F. 1979 The Semantic Organization of Some Simple Nouns and Verbs. *Journal of Memory and Language* 18:141–162.

IBM 1985 VM Programming in Logic (VM/PROLOG), Program Description and Operations Manual, IBM Corporation.

Ioup, G. 1975 Some Universals for Quantifier Scope. In: J. Kimball (Ed.), *Syntax and Semantics* 4. Academic Press, New York, NY.

Jackendoff, R. 1985 *Semantics and Cognition*. MIT Press, Cambridge, MA.

Johnson-Laird, P. N. 1983 *Mental Models*. Harvard University Press, Cambridge, MA.

Kamp, H. 1981 A Theory of Truth and Semantic Representation. In: J. Groenendijk, Th. Janssen, and M. Stokhof, Eds., *Formal Methods in the Study of Language*. Mathematisch Centrum, Amsterdam: 277–322.

Karttunen, L. 1971 "Possible and Must." *Syntax and Semantics* 1, 1–20. Academic Press, New York, NY.

Katz, J.J. 1972 *Semantic Theory*. Harper & Row, New York, NY.

Keil, F. C. 1979 *Semantic and Conceptual Development*, Harvard University Press, Cambridge, MA.

Labov, W. 1973 The Boundaries of Words and Their Meanings. In: C-J. N. Bailey and R. Shuy (Eds.), *New Ways of Analyzing Variation in English*. Georgetown University Press, Washington, D.C.

Lakoff, G. 1985 *Women, Fire and Other Dangerous Things*. University of Chicago Press, Chicago, IL.

Lloyd, J.W. 1984 *Foundations of Logic Programming*. Springer-Verlag, New York, NY.

Mann, W. and Thompson, S. 1987 Rhetorical Structure Theory: A Theory of Text Organization. ISI Reprint Series: ISI-RS-87-190. USC Information Sciences Institute. Marina del Rey, California 90292.

McCord, M. 1985 *The Lexical Base for Semantic Interpretation in a Prolog Parser*. Presented at workshop on the Lexicon, Parsing and Semantic Interpretation. City University of New York Graduate Center.

McCord, M. 1987 Natural Language Processing in Prolog. In: A. Walker, Ed., *Knowledge Systems and Prolog*. Addison-Wesley, Reading, MA.

McDowell, J. 1987 Assertion and Modality. Ph.D. dissertation. University of Southern California, Los Angeles, California 90089.

McDowell J. and Dahlgren, K. 1987 Commonsense Reasoning with Verbs. *Proc. IJCAI*.

Miller, G. 1978 Practical and Lexical Knowledge. In: E. Rosch and B. B. Lloyd (Eds.), *Cognition and Categorization*. Erlbaum, New York, NY.

Moens, M. and Steedman, M. 1987 Temporal Ontology in Natural Language. *Proceedings of the Association for Computational Linguistics*; 1–7.

Partee, B. 1984 Nominal and Temporal Anaphora. *Linguistics and Philosophy* 7:243–286.

Putnam, H. 1975 The Meaning of 'Meaning'. *Mind, Language and Reality*, Cambridge University Press, Cambridge, England.

Reichman, R. 1985 *Getting Computers to Talk Like You and Me*. MIT Press, Cambridge, MA.

Reinhart, T. 1982 Principles of Gestalt Perception in the Temporal Organization of Narrative Texts. Manuscript.

Reiter, R. 1980 A Logic for Default Reasoning. *Artificial Intelligence* 13:81–132.

Rickheit, G. and Strohner, H. 1985 *Inferences in Text Processing*. North-Holland, Amsterdam.

Rosch, E. 1975 Cognitive Representations of Semantic Categories. *Journal of Experimental Psychology-General* 204:192–233.

Rosch, E. and Mervis, C. B. 1975 Family Resemblances. *Cognitive Psychology* 7:573–605.

Rosch, E.; Mervis, C. B.; Gray, W. D.; Johnson, D. M.; and Boyes-Braem, P. 1976 Basic Objects in Natural Categories. *Cognitive Psychology* 8:382–439.

Schank, R. C. and Abelson, R. P. 1977 *Scripts, Plans, Goals and Understanding*. Erlbaum, Hillsdale, NJ.

Schubert, L. K.; Goebel, R. G.; and Cercone, N. J. 1979 The Structure and Organization of a Semantic Net for Comprehension and Inference. In: N.V. Findler (Ed.), *Associative Networks*. Academic Press, New York, NY.

Searle, J. and Vanderveken, D. 1985 *Foundations of Illocutionary Logic*. Cambridge University Press, Cambridge, England.

Smith, E. E. and Medin, D. L. 1981 *Categories and Concepts*, Harvard University Press, Cambridge, MA.

Steedman, M. 1977 Verbs, Time, and Modality. *Cognitive Science* 1:216–234.

Sterling, L. and Shapiro, E. 1986 *The Art of Prolog: Advanced Programming Techniques*. MIT Press, Cambridge, MA.

Stickel, M. 1986 A Prolog technology theorem prover: implementation by an extended prolog compiler. *Proceedings of the 8th International Conference on Automated Deduction*, J. Siekmann (ed.) *Lecture Notes in Computer Science Volume 230*. Springer-Verlag, New York, NY.

Strawson, P. C. 1953 *Individuals*. Methuen, London, England.

Trabasso, T. and Sperry, L. L. 1985 Causal Relatedness and Importance of Story Events. *Journal of Memory and Language* 24.1: 595–611.

Trabasso, T. and van den Broek, P. 1985 Causal Thinking and the Representation of Narrative Events. *Journal of Memory and Language* 24.5:612–630.

Vendler, Z. 1967 *Linguistics in Philosophy*. Cornell University Press, Ithaca, NY.

Wada, H. and Asher, N. 1986 BUILDRS: An Implementation of DR Theory and LFG. *Proceedings International Conference on Computational Linguistics* 540–545.

Webber, B. L. 1985 Discourse Model Synthesis. In M. Brady and R. Berwick, Eds., *Computational Models of Discourse*. MIT Press, Cambridge, MA, 267–330.

Notes

1. Dahlgren and McDowell are the main investigators. Stabler contributed the first-order logic and problem solver.
2. Nominalized verbs are treated as verbal concepts.
3. We are extremely grateful to Hajime Wada who wrote the original version of the DRT module to provide DRSs for a wide range of syntactic constructions. The present DRT module is an extension of his work. The lexical entries that form the NS data bases are the result of the careful, diligent efforts of Carol Lord, Robert Hagiwara, and Susan Mordechay. Susan Hirsh contributed the programs that generate the FOL data bases and performed other programming tasks.
4. The atoms *subj*, *obj*, *obliq*, *pobj*, which appear in the generic representations, signal which element of the sentence is to be accessed for the output response. In answering questions, the generic representations are mapped to a small set of canonical sentences whose slots are filled with elements from the input query. Consider the feature representation *implies(merchandise (obj))*, which is part of the generic entry for *buy*. If the query *What is implied if a man buys a truck?* is processed, since the direct object of *buy* in the query is *truck*, the response is *the truck is merchandise*. If the query had been *What is implied if John buys a house?*, the response would be *the house is merchandise*.
5. The analysis work was originally reported in Dahlgren and McDowell (1986b). We review that here and also report on the implementation.
6. For the most part, examples given in this paper are modified versions of actual sentences from the WSJ corpus that the KT researchers are using.
7. For readability, the remaining generic entries will be shown in English paraphrase rather than as they are coded.
8. The question arises as to how natural or likely such sentences would be in use. For example, would we be more likely to encounter *a letter from all the lawyers* rather than *a letter from every lawyer*. The standard answer, which we adopt, is that semantics is not a predictive theory. We can't tell what a person will say, but we have to be able to interpret whatever is said. We can restrict ourselves to likely expressions, but then we are putting ourselves in the position of predicting what is likely, and we might be wrong. We prefer to try to interpret what is possible, even if unlikely.
9. A full theoretical discussion of the issues involved with modals can be found in McDowell (1987).
10. The main predicate of a clause, whether it be a verb or a term that is the complement of the copula, carries a DR-theoretic event-type reference marker as its tense argument. Other predicates carry an argument which is linked to tense, but which is not a reference marker in DR-theoretic terms.
11. Actually, sentences interpreted as generics in English have a complex combination of the following features: indefinite NPs, present tense, copula, and/or inherently stative verb (but not the progressive). Not all of these features are always present in all generically interpreted sentences.
12. In this and the following sections we report on work in progress that is not yet fully implemented.