

AD-A085 893

ILLINOIS UNIV AT URBANA-CHAMPAIGN COORDINATED SCIENCE LAB F/6 3/7
THE SEMANTIC INTERPRETATION ON NOUN-NOUN MODIFICATION BY COMPUT--ETC(U)
JUN 79 T W FININ N00014-75-C-0612

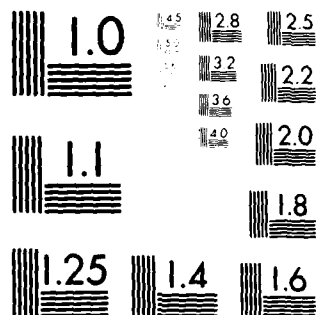
UNCLASSIFIED

WP-21

NL

$$\Delta Q = 0 \pm 9.3$$

END
DATE
FILMED
8-80
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ADA 085893

LEVEL II

12

6 The Semantic Interpretation of Noun-Noun Modification ^{by} Computer.

Working Paper/21

10 Timothy W. Finin

Advanced Automation Group
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801

1416 June 1979

12 431

Abstract

DTIC
ELECTE
S JUN 24 1980 D
E

This paper is a proposal for a Ph.D. thesis concerned with the semantic interpretation of noun-noun modification. The study of this topic is being coordinated with the development of the JETS natural language question answering system. One of the goals of this research is a computer program which will interpret instances of noun-noun modification within the domain of discourse of JETS. Related work on noun-noun modification in the disciplines of Linguistics and Artificial Intelligence is described and contrasted to the proposed research. The basic approach being take is described and the scope of the work is outlined.

Key Words and Phrases

natural language understanding, semantics, semantic interpretation, nominal compounds, nominalization, noun-noun modification, modification, linguistics, noun phrase

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

This work is supported by the Office of Naval Research under Contract N00014-75-C-0612.

Working papers are informal papers intended for internal use.

80 5 5 077

DDC FILE COPY

Finin

The Semantic Interpretation of
Noun-Noun Modification by Computer

A PhD Thesis Proposal

Timothy W. Finin

May 7, 1979

Accession For	
NTIS GRA&I	☒
DDC TAB	☐
Unannounced	☐
Justification	<i>per</i>
<i>Letter of File</i>	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or special
<i>A</i>	

committee

David L. Waltz (thesis advisor)
Geneva Belford
Jane Liu
Andrew Ortony
Sylvian Ray

This work is supported by the Office of Naval Research under contract
N00014-75-C-0612.

TABLE OF CONTENTS

I	The Topic	.1
I.1	The Context of the Research	.2
I.2	The Representational System	.3
I.3	Noun Modification	.3
I.4	Why Noun-noun Modification is a Difficult Problem	.4
I.5	Semantic Closure	.5
II	Related Work	16
II.1	Levi	16
II.2	Rhyne	22
II.3	Borgida	25
II.4	Marcus	26
III	The Approach	31
III.1	Rule Representation	31
III.2	Classes of Rules	31
IV	Scope of the proposed work	33
IV.1	Theoretical Issues and Problems.	33
IV.2	Implementational Issues	35
IV.3	Peripheral Issues	36
V	References	37

LIST OF FIGURES

FIGURE	Page
1. engine housing acid damage	7
2. A discourse context	10
3. A Meta-rule for Sets	11
4. A rule for plane - serial number	13
5. Another Meta-rule for Sets	14
6. Levi's Recoverably Deletable Predicates	20
7. The number of constrained vs. unconstrained trees	29
8. Examples which violate the Marcus Constraint	30

I

The Topic

I propose to study the semantics of certain types of basic noun modification with a focus on the modification of one noun by another. Under "basic noun modification" I include the prenominal modifiers, i.e. modification of a noun by adjectives, other nouns, possessives and prepositional phrases and the formation of compound nouns. I may to have something to say, but do not plan to focus on, the modification of nouns by quantifiers (e.g. some, any), articles (e.g. a, the) or relative clauses.

A major result of this research will be a computer program which will take instances of noun phrases and build a semantic representation which captures the intended meaning. This program will be designed and implemented as a component of a natural language question answering system which is being constructed concurrently by myself and others [FININ79].

I believe that this topic is a good vehicle for advancing the understanding of the semantics of natural language for the following reasons.

- (1) Noun modification is rich and productive.

It should be a good focus to bring out many of the general issues. The English language allows its speakers great freedom in the ways in which nouns can be modified. For example, I suggest that any two nouns can be related through modification in an appropriate context.
(1)

- (2) Noun modification has received little attention to date.

Most work on semantic interpretation in the computational linguistics field has focused on understanding the semantics of verbs and their case roles. In particular, the interpretation of noun-noun modification has been recognized as a thorny issue and generally put aside [WOODS72][BORGIDA].

- (3) Noun modification is essential.

The understanding of the semantics of simple noun modification will be important in almost any system which attempts to communicate with natural language. The rules, heuristics, and procedures developed in

¹ This makes a good game. Player A picks two nouns and player B tries to invent a context in which the two nouns make sense as a compound. This is reminiscent of the game invented by MIT linguists in which one tried to invent a context forcing the violation of any proposed selectional restriction.

the course of this work can have an immediate and practical application in existing and future natural language understanding systems [WALTZ76][CODD].

I.1 The Context of the Research

The goal of this research is the design of a system which will interpret the meaning of instances of simple noun modification. This system will be a component of the natural language data base accessing system JETS [FININ79]. The JETS system is currently being designed at the Coordinated Science Laboratory of the University of Illinois at Urbana-Champaign and is an outgrowth of our earlier system, PLANES [WALTZ76].

PLANES was developed to study the problem of natural language access to a large data base. The primary goal was to construct a system which would allow a non-programmer to obtain information from the data base by entering queries in a relatively unconstrained subset of English. Briefly, PLANES (1) received a request from the user; (2) parsed the request into an internal representation with a semantic grammar; (3) translated this representation into a formal query via a query generator; (4) executed the resulting query to retrieve the information; and (5) displayed this information in an appropriate manner.

The PLANES system accesses a large relational data base which contains information on Naval aircraft maintenance and flight records [NALDA][TENNAN78]. Maintenance records include such information as time and duration of the maintenance action, who performed it, what action was taken, and whether the service was scheduled or unscheduled. Aircraft flight data includes information such as the number of flights made by an aircraft for each day and month, the purpose of each flight, and the number and types of landings made for each flight. JETS is being designed to access the same data base.

The major thrust of our new design is to increase the coverage of the natural language processing. We see two nearly independent components to the concept of coverage. The conceptual coverage of a natural language system refers to the set of concepts that it can deal with. The linguistic coverage of a system refers to the linguistic knowledge that it has which enables it to understand variations in the way concepts are introduced, referenced, and described. A detailed discussion of coverage and the ways in which it can be measured can be found in [TENNAN79].

This research on noun modification addresses both component of coverage. Extending the linguistic coverage is the direct goal of this work. The semantic representation that JETS builds should not be highly sensitive to stylistic variations in the way a concept is described. For example, we want to build similar representations for the following phrases:

engine housing acid damage
acid damage to engine housings
acid damage to the housings of engines

damage by acid to engine housings
damage resulting from the corrosion of engine housings by acid
corrosion on engine housings
engine housing acid corrosion damage

This requires that the semantic interpretation rules be able to discover or infer the concepts which the words in the phrase refer to and the underlying relationships between them.

The indirect result of this work is the extension of the conceptual coverage of JETS. To compute similar semantic representations for the phrases in the above example, the concepts for engine, housing, acid, corrosion and damage must be represented in such a way as to enable and facilitate the discovery of the relationships between these words and words which they modify or are modified by. For example, the concept of damage should include the information that damage can be the result of another event (e.g. corrosion) and that the object that was damaged is important. We must represent the fact that acid can cause corrosion and that the corrosion event can lead to a state in which something is damaged.

I.2 The Representational System

How we have chosen to represent concepts in JETS is an important part of this research. The representational technology we are using is based on the frames paradigm [MINSKY][BOBROW][BRACHMAN]. A frame, as we are using it, is the basic unit of representation. Associated with a frame is a set of named slots which correspond to attributes of the entity being represented. Slots can contain values, of course, but can also contain such things as default values to be used when there are no explicit values, requirements which must be met before a value can be added, and procedures which are automatically invoked when a value is added, removed or accessed.

Individual frames are organized into an abstraction hierarchy which can be thought of as a directed tree of frames rooted at the most general concept (in our system, a THING). A particular frame can inherit attributes and values from its ancestors in the hierarchy as well as having its own information. This organization and associated inheritance is the most useful technique for capturing regularities and generalities in the concepts being represented.

I.3 Noun Modification

The linguistic process of noun modification is a central one in developing a theory of procedural semantics. In the English language, nouns can be modified in a variety of ways. For example, a noun can be modified by:

Articles	THE man
Adjectives	the TALL man

Nouns	the REPAIR man
Clauses	the man WHO LEFT
Genitives	the FBI's man
Prep. Phrases	the man FROM MARS
Ordinals	the THREE men
Quantifiers	EVERY man
Participles	the APPROACHING man
	the DEFEATED man

In the field of Computational Linguistics, a great deal of attention has been directed toward developing semantic theories at the level of the clause. One reason for this is the existence of an elegant paradigm: case frame theory. This paradigm, first proposed by Fillmore [FILLM068], has been adapted and used in almost all AI natural language research. Central to this approach is the fact that the semantic interpretation of clauses is strongly governed by the main verb. With each verb, (or verb-sense if the word allows multiple senses) one can associate data structures and procedures which guide the interpretation of syntactic constituents found with it. For example, the verb GIVE can have associated with it such case roles as an AGENT, RECIPIENT, OBJECT, TIME, MANNER, INSTRUMENT, etc. Determining the semantic relationship (i.e. case role) between the verb GIVE and its syntactically associated constituents involves using a variety of syntactic, semantic and pragmatic clues [FININ76].

I would like to develop a similar theory or paradigm for modification at the noun phrase level. The interaction of a verb and its case role fillers is, in many ways, similar to the interaction between a noun and its modifiers. In both cases the meaning is most strongly determined by the head of the structure - the verb in the clause and the head noun in the noun phrase. The modifiers (case role fillers in the clause and prenominal modifiers in the noun phrase) typically add additional information to a structure called forth by the head. In both cases the possible interpretations of the modifying words can affect the selection of the correct sense of the modified word if it has more than one sense.

1.4 Why Noun-noun Modification is a Difficult Problem

The basic task of computing the semantic interpretation of noun-noun modification is easy to state. Given two nouns in which the first modifies the second, we need to discover the relationship which the speaker intends to hold between them. For example, in "aircraft engine" the relationship might be part of (the engine is part of an aircraft) and in "meeting room" it might be location of (the room in which a meeting takes place).

This is a very general task. One can view the interpretation of other English structures in this same way. At the clause level, we might view the interpretation problem as one of discovering the relationship between the main verb and its subject, objects and adverbial modifiers. At the noun phrase level we can view interpretation as a task of relating the head noun to its

prenominal modifiers, prepositional phrases and relative clauses. Is there anything different about noun-noun modification which warrants a special study? My answer, not surprisingly, is yes. The distinguishing feature of noun-noun modification is that it is an open form of modification (a notion I will define momentarily) in which there are no syntactic or structural clues which guide one to the intended interpretations.

By an open form of modification I mean one in which the participating constituents (in this case the two nouns) are both chosen from an open syntactic category. This is in contrast to such closed forms as modification of a noun by a article or the modification of a verb by tense and aspect morphemes. If we want to formalize what it means for a noun to be modified by a article, we need only formalize what it means for a noun to be modified by the words a, an and the. This is not to say that the problem is trivial for such closed forms, only that it is from the start of a much lower order of difficulty.⁽²⁾

The second source of difficulty in the semantic interpretation of noun-noun modification is the lack of additional syntactic clues to the meaning. All we are given are the two nouns and the hypothesis that the first modifies the second. In the interpretation of a clause, one has several syntactic clues with which to work. Word order is the most obvious. Other clues include the presence of particles and the marking of some case roles by particular prepositions. One can view the interpretation of prepositional phrases as refining the relationship, determined by the preposition, between two noun phrases. Here, at least, the preposition suggests a set of relationships that might hold between the two noun phrases.

I.5 Semantic Closure

A major part of our work on the JETS system is centered on achieving a high degree of closure. By closure, we mean the ability to handle user utterances which are consistent with the conceptual domain of the system, but which have not been foreseen. Woods [WOODS77] defines the concept of closure for natural language processors as:

"The difficulty in natural language understanding is not so much being able to formulate rules for handling phenomena exhibited in a particular dialog, but to do so in such a way that closure is eventually obtained -- i.e., subsequent instances of the same or similar phenomena will not require additional or different rules,

² An example of a difficulty in the interpretation of the apparently simple case of the article the is the fact that "the X" where X is singular might be an instance of reference to a definite X known to both the speaker and hearer or it might be a reference to the generic class of X's. Contrast the sentences "the computer is down this afternoon" and "the computer is a useful device"

but will be handled automatically by generalized rules. The formulation of such rules requires a good formalism for expressing general rules and a methodology for obtaining the correct degree of interdependence among individual rules. It also requires a good linguistic intuition and/or knowledge of linguistic results for determining the correct generalizations of the phenomena."

A goal of my work, then, is to achieve a high degree of closure in the semantic interpretation of simple noun modification and, in particular, noun-noun modification.

The problem of interpreting strings of nouns related through modification is a complex one. As a first order theory, I divide the problem into three subproblems: lexical interpretation, modifier parsing and concept modification.

By lexical interpretation I mean the process of mapping the lexical items (in this case the nouns in the string) into appropriate concepts. The principal difficulty here is handling words with multiple senses.

Modifier parsing is the process of discovering the internal structure associated with the string of nouns or the concepts which result after lexical interpretation. For example, a string of three nouns, N1 N2 N3, might have the structure ((N1 N2) N3) or the structure (N1 (N2 N3)). The first structure would be chosen for the string "engine damage reports" and the second for the string "replacement oil pump".

The term conceptual modification refers to the problem of assigning an interpretation to an instance of one concept modifying another concept. For example, when the ENGINE concept modifies the DAMAGE concept in the phrase "engine damage" we want to fill the damaged object role in the DAMAGE concept with the ENGINE concept. A more complex example is the interpretation of the phrase "engine housing acid damage". Here, the desired result is a something like the network of frames shown in figure 1. A prose description of this network would be:

A RESULT of a DAMAGE event in which the damaged object is the HOUSING part of an ENGINE and the cause is a CORROSION event in which the corrosive agent is an ACID and the corroded object is the HOUSING part of an ENGINE.

These three subproblems are, of course, interrelated and cannot be completely decoupled. In my initial research I am concentrating on the problem of conceptual modification. The goal is that, given any two concepts that are correctly interpreted by the system, their combination (i.e. through modification) will be correctly interpreted by the system. Note that the correct interpretation of a concept does not imply that the system should be able to "handle" the concept in the sense of answering questions about the concept or relating it to the data base.

The problem of interpreting noun-noun modification brings the issue of closure into focus. The essential feature of noun-noun modification is that

RESULT2

...
event = DAMAGE48
...

DAMAGE48

...
cause = CORROSION16
damage agent = ACID3
damaged object = HOUSING12
...

CORRSOSION16

...
corrosive agent = ACID3
corroded object = HOUSING12
result = RESULT2
...

ACID3

...

HOUSING12

...
part of = ENGINE6
...

ENGINE6

...

engine housing acid damage

figure 1

May 7 1979

the semantic relationship which exists between the two nouns is not explicit in the utterance. Moreover, a large number of relationships may, in principle, be possible between the two concepts represented by the nouns. It is the responsibility of the system to attempt to infer or discover an appropriate relationship, given its understanding of the two concepts involved, general pragmatic knowledge, and the current discourse context.

An example: Time

As an example, consider the use of a time phrase used to modify a noun, as in the phrases:

January Skyhawks repairs
1976 flights

If the system can interpret phrases referring to time (as almost any system must) then it should attempt to interpret the modification of any other concept which could conceivably have a time phrase attached to it.

In the semantics we are developing for JETS, a time phrase can only be used to modify a concept which is, or can be viewed as, a kind of an EVENT. A minimal amount of closure is achieved when any event or event related concept can be successfully modified by a TIME concept. What if the modified concept is not an EVENT but something else, say an OBJECT? If a time phrase is hypothesized to modify something which is a kind of OBJECT, we want the system to attempt to derive an underlying event associated with that object to attach the time phrase to. For example, in the standard PARTS-SUPPLIERS-PROJECTS (3) domain, the phrase "January parts" might suggest the interpretations:

parts which were shipped in January
parts which were received in January
parts which were ordered in January

In such an impoverished domain this is almost trivial, as one can precompute the set of events in which a concept can partake.

In a semantically rich domain, such as our 3-M data base [NALDA], the problem is much more difficult. One can not (or perhaps should not) always enumerate the potential relationships which might exist between even two simple concepts. The ability to handle references to entities and relations mentioned earlier in the discourse makes the problem even more complex. This allows for more potential relationships between any two concepts. For example, the phrase

3 This domain is often used to describe the operations of data base query systems. Codd's RENDEZVOUS system is one natural language system which uses this domain. In its simplest form, it contains information on parts (e.g. part numbers and names), projects (e.g. name, location, inventory of parts), suppliers (e.g. name, location, rating), and shipments (e.g. from which supplier, to what project, part number, quantity).

the January planes

could be used to refer to a set of planes introduced previously in the discourse. The successful interpretation of this phrase would require a search through the recent discourse to discover a set of planes which was involved in an event which occurred in January. For example, the context might have been the one shown in figure 2. In this case, "the January planes" should be interpreted as referring to "the planes which received engine maintenance in January 1978".

An example: Sets

We have introduced our JETS system to the concept of a SET. In order to achieve a high degree of semantic closure, the system should be able to form the concept of a SET over a wide domain of objects. Our previous system, PLANES, handled sets in an unsatisfactory way. One could refer to sets of objects of certain types but not others. For example, PLANES could understand descriptions of and references to a set of aircraft or maintenance codes but it could not handle sets of parts or "how malfunctioned" codes. Such shortcomings are particularly bad in that they mislead users. If a user was successful in using a description of a set of objects which PLANES understood, he could quite reasonably infer that PLANES understood the general concept of a set and could form one of arbitrary objects.

Given that sets can be represented and formed in a uniform way over the widest possible domain, we must turn our attention to issues of interpreting the modification of the set concept. The ability to form sets of arbitrary elements will be of limited use if the semantic interpretation rules do not allow one to modify such sets in a general way. Thus, if the system knows what it means for concept X to modify concept Y, then it should know what it means for concept X to modify a concept Z where Z is a set whose members are concepts which are Y's.

My approach is to include a meta-rule for sets which uses rules applicable to the particular domain of a set. My SET frame has a slot for a typical member as well as one to receive the actual members, if there are any. The typical member slot refers to a frame which describes the typical member of the set. Whenever we wish to modify a set by another concept, this meta-rule will search for primitive rules which interpret modification of the set's typical member by that concept. The rules which are found to be applicable are then invoked on the typical member and to each of the set's individual members, if any exist. This meta-rule for sets is shown in figure 3.

Let's examine the use of this meta-rule for sets in the interpretation of the phrase:

Planes 3, 5, and 48

At the concept level, this phrase is represented as the general PLANE concept modifying a SET concept. This particular SET has the INTEGER concept for its typical member and, as individual members, the concepts for the integer 3, the integer 5, and the integer 48.

<user> Show me the engine maintenance performed on F4's in the last three months.

<JETS>	<u>DATE</u>	<u>PLANE</u>	<u>MAINTENANCE CODE</u>
	1/2/78	3	...
	1/10/78	3	...
	1/12/78	23	...
	1/20/78	23	...
	1/26/78	3	...
	1/28/78	48	...
	2/6/78	4	...
	2/10/78	32	...
	2/10/78	23	...
	.	.	.
	.	.	.
	.	.	.

<user> Which of the January planes also required maintenance in December?

A discourse context

figure 2

if <concept> modifies a <set> then:

find the typical member of the <set>.

find an applicable rule which interprets the
modification of typical member by <concept>.

invoke the rule on the typical member.

invoke the rule on each of the members of the
<set>.

return the newly modified <set>.

A Meta-rule for Sets

figure 3

One of the rules applicable when a PLANE modifies an INTEGER, is a rule which interprets the integer as representing the plane's serial number (see figure 4. In our world, the only planes which have serial numbers are those in the data base. These planes are represented by the more specific concept 3M-PLANE. The action of this rule is to view (4) this plane concept as a 3M-PLANE and the integer concept as a SERIAL NUMBER. The 3M-PLANE's serial number slot is then filled with the SERIAL-NUMBER concept.

When this rule is found by the meta-rule for sets, it is applied to both the set's typical member filler (in this case the generic INTEGER concept) and to the set's members (which are the integers 3, 5 and 48). The result of all this is described by the concept frame:

A SET with

typical member = a 3M-PLANE

members = a 3M-PLANE with

serial number = an INTEGER with value = 3

a 3M-PLANE with

serial number = an INTEGER with value = 5

a 3M-PLANE with

serial number = an INTEGER with value = 48

To handle the case where a SET is used to modify another concept, we include another meta-rule, shown in figure 5. Consider the role of this rule in the interpretation of the phrase:

radar and navigation equipment failures

The phrase "radar and navigation equipment" is interpreted as a SET whose typical member is a FUNCTIONAL-SUBSYSTEM and whose members are the concepts RADAR-SUBSYSTEM and NAVIGATION-SUBSYSTEM. Or, expressed in a simple frame description language:

a SET with

typical member = a FUNCTIONAL-SUBSYSTEM

members = a RADAR-SUBSYSTEM

a NAVIGATION-SUBSYSTEM

Note that the rule which formed this instance of the set concept characterizes the typical member as a functional subsystem. The heuristic used finds the

4 Viewing a concept X as a concept Y is a process which maps the information in X into a newly instantiated Y concept. If X is a kind of Y, no mapping need be done, however.

if (a plane) modifies (an integer) then
 view (the integer) as (a serial number).
 view (the plane) as (a 3m-plane).
 put (the serial-number) in (the 3m-plane)'s
 serial number slot.
 return (the 3m-plane).

A rule for plane - serial number

figure 4

if a <set> modifies a <concept> then:

find the typical member of the <set>.

find a rule which interprets the modification of the <concept> by the typical member.

form a new instantiation of a SET.

fill the new typical member slot by invoking the rule on the typical member and <concept>.

fill the members of the new SET by invoking the rule on the member of the old SET and the old <concept>.

Return the new SET.

Another Meta-rule for Sets

figure 5

least general concept in the abstraction hierarchy to which all of the members belong and uses this as a description of the typical member. This is not the only possibility, of course. One could just as well use a much narrower generalization. In this case the candidate would be "either a radar-subsystem or a navigation subsystem".

In interpreting the modification of the FAILURE concept by this SET, the meta-rule for sets is invoked and attempts to find rules which guide the interpretation of a kind of FUNCTIONAL-SUBSYSTEM modifying FAILURE. The rule which is most applicable is one which interprets the modifying concept (the subsystem) as filling the failure location role in the FAILURE concept. The final interpretation of this phrase results in a SET of FAILURES in which the typical member is a FAILURE in a FUNCTIONAL-SUBSYSTEM and which contains two members: a FAILURE in the RADAR-SUBSYSTEM and a FAILURE in the NAVIGATION-SUBSYSTEM. In other words:

a SET with

typical member = a FAILURE with
failure location = a FUNCTIONAL-SUBSYSTEM

members = a FAILURE with failure location = RADAR-SUBSYSTEM
a FAILURE with failure location = NAVIGATION-SUBSYSTEM

II

Related Work

This section describes previous work which relates to the problem of interpreting the meaning of instances of simple noun modification. There are two primary sources of related research: Linguistics and Computational Linguistics/Artificial Intelligence. Many linguists have been drawn to the problem of the formation of nominal compounds. Primarily, their studies have been concerned with the problem of discovering the constraints our language places on the formation of compounds rather than the problem of interpreting their meaning. AI research, on the other hand, is primarily interested in computing a meaning for an instance of modification, be it nominal or other. Relevant AI research can be found in work on natural language understanding and question answering systems, of course. The work in the more general area of knowledge understanding and representation is also relevant, as the interpretation of noun modification requires a good knowledge representation system.

The rest of this section is devoted to a more detailed description of the work by one linguist (Judith Levi) and several AI researchers.

II.1 Levi

Judith Levi [LEVI] has made an extensive exploration of one set of prenominal modifiers. She defines the notion of a Complex Nominal (CN) which includes three kinds of expressions, all of which are fundamentally instances of noun-noun modification.

The first set, nominal compounds, covers the classical noun-noun modification form. Nominal compounds include instances in which the prenominal modifier is also a noun. Examples are:

apple cake
brush fire
dog house
piston ring
wing tip

The second set, nominalizations, include pairs in which the modifying word is a noun and the modified head noun is derived from an underlying verb. Example from this set are:

city planner
data encoding
signal detection
engine damage

maintenance reports

The third set defined by Levi as "NP's with non-predicating adjectives". This set includes phrases in which the head noun is modified by a word having the surface form of an adjective, but an underlying derivation from a noun. Examples are:

rural rout
electrical engineer
musical clock
constitutional amendment

Levi gives many detailed arguments to support the position of viewing these "pseudo-adjectives" as noun-like terms. Because I believe this analysis to be valid and relevant to my own work, I will sketch some of the principal arguments here.

Non-predicating adjectives (NPAs) do not normally appear in the predicate, or post-copula, position where a bona fide adjective can be used. To see this, consider the potential paraphrases listed below.

a chemical engineer	an engineer who is chemical
an atomic bomb	a bomb which is atomic
a linguistic argument	an argument which is linguistic

More appropriate paraphrases might be "an engineer whose specialty is chemistry", "a bomb in which the explosive energy comes from an atom", and "an argument about Linguistics". There are some NPAs which can also serve as predicating adjectives, but these are not synonymous when used as such, as the following phrase pairs show:

a criminal lawyer	a lawyer who is criminal
a logical fallacy	a fallacy which is logical
dramatic criticism	criticism which is dramatic

Similarly, we can distinguish between predicating and non-predicating adjectives by their ability to be modified by "degree adverbials" such as very, quite, and slightly. For example, we do not say "a very atomic bomb" or "a very electrical conductor". For NPAs which have a predicating sense, the use of a degree adverbial forces the selection of the predicating reading. Thus one might refer to R. M. Nixon as a "very criminal lawyer".

There is also evidence which supports the claim that NPAs are derived from ancestor nouns. Briefly these are:

- (1) NPAs can be conjoined with nouns.

Generally, English only allows like constituents to be conjoined. We are able to conjoin a noun and NPA, as in "solar and gas heating systems" and "domestic and farm animals".

- (2) NPAs can be categorized by semantic features assigned to nouns.

One can assign such semantic features as DEFINITE, ANIMATE and GENDER to NPAs as well as to nominals. For example, presidential and feline would be +ANIMATE whereas electric and automotive would be -ANIMATE.

(3) NPAs are amenable to a case-frame analysis.

Unlike bona fide adjectives, NPAs are readily analyzed in terms of case relations, particularly when the modified noun is a nominalized verb. For example, presidential can be seen as filling the agentive case in "presidential veto" and lunar as filling the objective case in "lunar explorations".

(4) NPAs may not be nominalized.

Unlike adjectives and like nouns, NPAs may not be nominalized. This is easiest to see when one considers adjectives which have both a predicating and a non-predicating reading. When an adjective is being used with its predicating sense it can be nominalized but not when it is being used with its non-predicating sense. Consider the word "mechanical" which has a predicating sense in "mechanical reaction" and a non-predicating sense in "mechanical engineer". When the predicating sense of this word is used it can be nominalized as in "the mechanicalness of the reaction". The non-predicating sense can not undergo nominalization, ruling out a phrase like "the mechanicalness of the engineer". Similar examples are listed below.

predicating sense	non-predicating sense
mechanical reaction The mechanicalness of her reaction	mechanical engineer *The mechanicalness of the engineer
nervous reaction the nervousness of the reaction	nervous disorder *the nervousness of the disorder
marginal contribution the marginality of the contribution	marginal width *the marginality of the width

Levi's analysis

A brief summary of her analysis is as follows. Complex nominals are derived from an underlying structure which consists of a head noun modified by a sentential construction. This construction can be either a relative clause or a NP complement. All CNs are derived through the application of one of two transformations: the deletion of the predicate in the modifying sentence, or the nominalization of the predicate in the modifying sentence.

CNs derived from Predicate Deletion

The predicate deletion transformation may only apply to a proposition whose predicate is a member of a small set of Recoverably Deletable Predicates (RDP). This set consists of the following eight predicates:

CAUSE	HAVE
BE	USE
FOR	IN
ABOUT	FROM

Figure 6 gives some examples, many from Levi, which exhibit the possibilities. Note that for three of the RDPs (CAUSE, HAVE and MAKE), there are two examples, corresponding to the active and passive forms of the deleted predicate.

Since the predicate relating the head noun and its modifier is deleted by this transformation, the meaning of the resulting CN is multiply ambiguous. Each of these predicates are good candidates for the relation between the two nominals. The ambiguity is constrained, however, by the smallness of the set of potential predicates, the RDP. In practice, this ambiguity is further reduced by semantic and pragmatic knowledge as well as the discourse context.

CNs derived from Nominalization

The second class of CNs discussed by Levi are those derived from the nominalization of the predicate in an underlying proposition modifying the head noun. Some example from this class are:

- faculty meeting
- tree traversal
- urban studies
- fire detection
- mathematics teacher

In these cases, the head noun is derived from an underlying verb (i.e. meet, traverse, study, detect, amplify and teach). The modifying noun can stand in one of two relationships to this underlying verb. In subjective nominalization, the modifying noun is derived from the subject of the verb in the underlying form. Examples are:

- University purchases
- acid corrosion
- bird damage

Objective nominalization occurs when the modifying noun is derived from the object relation of the verb in the underlying sentence. Examples from this case are:

- automobile purchases
- engine corrosion
- propeller damage

A third case, which Levi analyzes as a sub-case of objective nominalizations, is one in which the head noun denotes the agent of an action rather than the action itself. Examples are:

- science teacher

RDP	phrase	paraphrase
CAUSE	tear gas drug deaths	gas which causes tears deaths which are caused by drugs
HAVE	gun boat box top	boat which has guns top that a box has
MAKE	musical clock program errors	clock which makes music errors made by a program
USE	steam iron	iron which uses steam
BE	soldier ant	ant which is a soldier
IN	city bus	bus which is in a city
FOR	animal doctor	doctor who is for animals
FROM	city visitors	visitors who are from a city
ABOUT	physics book	book which is about physics

Levi's Recoverably Deletable Predicates

figure 6

coal miner
electrical conductor

In these cases, the modifying noun or non-predicating adjective is again the object of the verb in the underlying form.

Critique

Levi's work is very impressive and covers much of the groundwork concerning noun-noun modification. It is also a wonderfully rich source of examples of such modification. The focus of her work, however, is quite different from that proposed here. The nature of this difference is a common one when comparing work done by Linguists and AI researchers. This section briefly discusses some of the major points at which her work and my own diverge.

Levi's analysis is primarily Syntactic. 1 The primary goal of the study was the elaboration of detailed derivations for complex nominals within the framework of generative semantics. Generative semantics, the name notwithstanding, is heavily concerned with the form and structure of language and the structural transformations which operate on sentences.

Levi's semantic analysis is too shallow. My primary objection is that her set of eight Recoverably Deletable Predicates are extremely vague. It is my feeling that such vague predicates as HAVE, FOR and IN should not be the stopping point of the semantic analysis. Most of the difficult but interesting work is in specifying exactly what these predicates mean for a particular case of noun noun modification. The following quote from Levi suggests that she would acknowledge this view:

"In the course of this exploration, it has become clear that a complete description of the role of complex nominals in natural language must include not only the kinds of syntactic and semantic facts that formal derivations can account for, but also a description of the broader semantic and pragmatic principals that influence the ways in which both speaker and hearers manipulate the formal regularities in actual discourse. Although this latter aspect of the grammar of CNs lies outside the scope of this study, its indisputable relevance to "the larger picture" as well as its intrinsic interest suggests that a brief discussion of the major issues may be appropriately included here."

Levi does not treat the use of complex nominals in definite descriptions. When one is making an anaphoric reference it is common to find complex nominals being used to "telescope" the description of the referent. An example described earlier in this paper showed how the phrase "the January planes" could refer to a set of planes which had engine maintenance in January. Levi's approach can not possibly help us for such cases.

Ambiguity is ignored. Levi is content to reduce the ambiguity to a small set of cases. The problem of using deeper semantic and pragmatic knowledge to resolve the ambiguity should, I believe, be a part of the analysis. Again, I feel that she would not quarrel with this. She says:

"What is not yet clear is how best to represent those kinds of knowledge that are not "strictly grammatical" in the common sense of the term (especially the basically ephemeral knowledge of "institutionalised" readings of CNs, and the highly context-sensitive variables which enter into our stylistic judgements), or how these different but related kinds of knowledge may best be integrated within a single description."

Processing strategies are ignored. Finally, her approach is more from the viewpoint of language production rather than language understanding. Little attention is directed to the problems of how people (or machines) might process complex nominals to extract an interpretation consistent with a larger context.

II.2 Rhyme

James Rhyme [RHYNE] has made a study of nominal compounding within a computational linguistic framework. In his work he developed a procedural model for generating nominal compounds from a noun phrase represented in a case-frame formalism. His basic analysis of nominal compounds is that their interpretation and generation depends on the existence of a characteristic relationship between the modifying noun and a verb in a paraphrase using a relative clause construction. Compounding is a process of systematically deleting information from an utterance just when the speaker expects the hearer to be able to reconstruct it.

Linguistic issues

Rhyme identifies three structural forms of nominal compounds in his work. The first (N-N) is one in which a noun is modified by another (surface) noun. The compounds computer terminal, telephone cord and aircraft engine are examples of this form. He briefly discusses the potential ambiguity which arises when there are two or more modifying nouns and notes that, in English, there is a slight preference for interpreting such phrases in a left-to-right manner. Thus, it is more common to find the (N-N)-N forms like:

typewriter repair man
electric typewriter repair man
engine damage report
jet engine damage report

than the N-(N-N) forms given below:

liquid roach poison
aluminum water pumps
January aircraft repairs

The second and third forms are N-participial-N and N-gerund-N, respectively. Rhyme discusses these forms only in passing.

Rhyne argues that nominal compounds in English are the result of one of two processes. The first involves the reduction of a relative clause followed by the preposing of the remaining element. The second process is one in which the verb contained in the relative clause is nominalized and then preposed to modify the head noun.

Rhyne chose for an underlying representation a shallow case grammar rather than a deep case representation. This was motivated by his belief that the rules used to generate nominal compounds are primarily lexical. That is, their relevance and application strongly depend on the actual lexical items (e.g. words) which appear in the case frames. His case grammar is fairly typical, being similar to others developed at the University of Texas [SIMMONS]. It includes the following verb case roles:

PERFORMER
CAUSE
ENABLER
OBJECT
GOAL
SOURCE
LOCATION
MEANS

In addition, he uses two "structural" case roles: RELCLS and COMP. The RELCLS (relative clause) role is used to attach a relative clause to a noun. The COMP (compound) role is used to attach a modifying compound to a noun.

Constraints

Rhyne proposes three general constraints on potential rules for generating nominal compounds. The first is that nominal compounds are used to express characteristic or habitual relationships. A shrimp boat is a boat which is characteristically used to catch shrimp. The fact that the same boat was once used to catch sharks does not allow one to refer to it as a shark boat.

The second constraint involves the use of a proper noun as a noun modifier. Rhyne claims that this can only occur when the proper noun is the name of a process or a source, performer or goal of an act of giving.

The third constraint involves the degree to which terms in his rules match the lexical items in the structure being transformed. Rhyne's rules include class terms which could potentially match many instantiations. For example, one rule might involve the class <person>, which could match the lexical items man, woman, child, Indian or midget. He states that compounds are not generally formed when the lexical item is several levels below the class term used in the rule. Thus, a rule could be given which transformed "a <person> who repairs things" into "a repair <person>". This would generate the compound "repair man" from "a man who repairs things" but would not transform "a midget who repairs things" into "a repair midget".

Rhyne's Computer Model

Rhyne developed a simple computer model which transformed expressions in his shallow case grammar into surface nominal compounds. It consisted of a recursive rule interpreter and a collection of lexical transformation rules. As an example of one of his rules, consider a rule to map "a market which sells flowers" into the compound "flower market". It might be expressed as:

```
(market (RELCLS (sell +CHARACTERISTIC
                (LOC market)
                (OBJ flowers))))
```

==>

```
(market (COMP flowers))
```

We can write a more productive version of this rule by replacing the term "flowers" with the generalization <goods>. The rule would then be:

```
(market (RELCLS (sell +CHARACTERISTIC
                (LOC market)
                (OBJ <goods>))))
```

==>

```
(market (COMP <goods>))
```

This rule would then account for the following compounds:

```
meat market
fish market
computer market
```

Critique

Rhyne approaches the general problem from the point of view of language production rather than language understanding. This places the focus on issues which are somewhat different from mine. Rhyne begins with a complete semantic representation of a phrase which contains all the relevant information. His goal is to produce a surface level representation of the phrase using nominal compounds whenever possible. This bypasses many of the problems which I hope to address in my research. In principle, the production rules he uses to generate the surface level phrase from the internal semantic representation could be reversed to produce the semantic representation from the surface one. However, one then has to face the problems of multiple word senses, ambiguity (both structural and semantic), and the interaction of intra- and inter-sentence context.

II.3 Borgida

Alexander Borgida's 1975 thesis [BORGIDA] contains a chapter on the semantic interpretation of the noun phrase in which he proposes a simple classification of noun-noun modification types. His basic approach is to find an underlying verb which relates the head noun and its modifier. Given a head noun N and a modifying noun M, his classification is as follows.

- (1) The head noun is an agental nominalization (e.g. owner, student, buyer).

In this case the underlying relationship between N and M is the event (or verb) from which N is derived. The head noun fills the agent case role of the event and the modifying noun can fill any of the other case roles. For example, the modifier M could fill the object role of the verb ("physics teacher"), the place/location role ("university student") or the time characteristic role ("night guard").

- (2) The head noun is a result nominal (e.g. application).

As in the first case, the underlying relationship is determined by the verb that N is derived from. In this case, the agent case role is also free to accept the modifying noun M, as in "student application".

- (3) The modifier is derived from a verb.

In this case the underlying relationship is determined by the verb from which the modifier is derived. The head noun can fill any of the case roles associated with this verbs case frame. Examples of this case are:

reception committee	(N is <u>agent</u>)
fish hook	(N is <u>instrument</u>)
completion date	(N is <u>time characteristic</u>)
meeting room	(N is <u>place/location</u>)

- (4) Neither the head noun nor the modifier is derived from a verb, but one of them is closely related to a verb.

In this vague class, Borgida puts such examples as "steel factory", "dog house" and "university degree". His idea is that a noun such as "factory" is closely associated with the verb "make". The interpretation of "steel factory" would then be something like "a factory in which steel is made".

- (5) A common fixed relationship holds between the head noun and its modifier.

This case includes a class of compounds which are related by a relation from a set of common fixed relationships. As examples of these relationships, Borgida gives: part of, type of, made of, and produces/yields.

In passing, Borgida mentions the problem on representing the limited productivity of noun compounding rules. One could formulate a rule to interpret "bus stop", "train stop" and (in general) "vehicle stop". Presumably, the interpretation would involve the case frame for the verb "stop" from which the noun "stop" is derived and result in "a place where a bus/train/vehicle stops". The compound "man stop" is perceived as bizarre, however, even when the same semantic relation ship holds (i.e. we wish to refer to a place where a person stops). His proposal is to indicate on some nodes of his semantic network the compounds they can form and then to deduce that all subconcepts of these node may also participate in like compounds.

critique

Borgida admits that his discussion of noun-noun modification is brief. In fact, he says of it that it "seems to be the most complicated form of noun modification". His analysis is too dependent on finding an underlying verb or event associated with one of the two nouns. This seems to be a reasonable heuristic, especially when one is dealing with nominalizations of verbs, but it does lead to several problems. One problem is that he slights the cases where it is not evident what the related verb/event might be (his cases number four and five). He offers no suggestion as to what it means for a verb to be closely related to a noun or how this is to be represented in a uniform way. Similarly, he gives no rules or heuristics for evaluating the appropriateness of relationships from the fixed class (his case number five).

A more serious problem arises when we attempt to constrain the productivity of some forms. Using his own example of a rule to interpret "bus stop", it seems we need some way to prevent the general find-a-related-verb heuristic from producing the interpretation of "man stop" as a place where people stop. Even if we can constrain the modifier to be a kind of vehicle, we run into trouble. Such a constraint would allow the bizarre compounds:

plane stop	golf-cart stop
van stop	fork-lift stop
motorcycle stop	

What is needed is a representation of a sense of the word stop which means:

A place where a vehicle stops to take on or let off things (especially people) for the purpose of transporting these things to another place.

II.4 Marcus

In his thesis, Marcus [MARCUS] proposes a simple theory to solve what I call the modifier parsing problem. His hypothesis is that a parser with a buffer "window" of three constituents is sufficient to analyze deterministically noun-noun modifier strings of arbitrary length. Given the phrases:

water meter adjustment screw
 ion thruster performance calibration
 boron epoxy rocket motor chambers

he wants to produce the parses:

```
[[[ water meter ] cover ][ adjustment screw ]]  

[[ ion thruster ][ performance calibration ]]  

[[ boron epoxy ][[ rocket motor ] chambers ]]
```

His procedure is based on two assumptions. First, he assumes a semantic component which can decide upon the relative "goodness" of two possible noun-noun modifier pairs. For example, given the pairs "water meter" and "meter cover", this oracle would judge the first to be superior to the second, even though both are acceptable. The second assumption (which is the one with theoretical interest) is that arbitrarily long strings of nouns can be analyzed by examining the three left-most nouns (simple or compounded) in the string.

A third assumption which he does not explicitly mention captures the slight bias in English for constructions like ((N N) N) over (N (N N)). The kernel of this algorithm is a rule for parsing a string of three nouns. Assume that there are three nouns in the buffer: N1, N2 and N3. Let [N1 N2] stand for the modification of N2 by N1. His rule is then:

If [N2 N3] is semantically better than [N1 N2] then replace the buffer with N1 [N2 N3]. Otherwise, replace the buffer with [N1 N2] N3.

These assumption yield a simple algorithm which reads the first three nouns of a long string into the buffer and forms a compound noun out of either the first and second noun or the second and third (under the direction of his semantic oracle). The buffer is then contracted and the next noun is pulled in. This process is repeated until the buffer has been reduced to two nouns, the first of which is taken to modify the second.

This is a highly interesting theory and one which would be important if true. His theory would greatly reduce the number of possible parses which would need to be considered for a long string of nouns. Without this constraint, the number of different parses for a string of N nouns is given by the recurrence relation:

$$f(n) = \sum_{i=1}^{n-1} [f(i) * f(n-i)]$$

which has the closed form:

$$f(n) = \frac{1}{n+1} * \binom{2n}{n}$$

This function is bounded from above by the inequality:

$$f(n) < \frac{n^4}{n \sqrt{\pi n}} + O(n^{-5/2})$$

With the Marcus constraint, the number of possible parses is reduced to:

$$f'(n) = 2^{n-2}$$

Figure 7 gives a table which shows values for f and f' for some small values of n .

Another way to state Marcus's constraint is to characterize the trees that can be produced. Let the right depth of a leaf of a tree be the number of right daughter links traversed on a path from the root of the tree to a that leaf. The maximum right depth of a tree is the maximum right depth over its leaves. Marcus's constraint is that the parse tree have a maximum right depth of two or less.

critique

I think that the Marcus constraint is very interesting in that it is true for the great majority of long strings of nouns related through modification. There are counter-examples, however. Marcus himself mentions a single counter example which he discovered. He assigns the phrase:

1970 balloon flight solar cell standardization program

the structure:

[1970 [[balloon flight][[solar-cell standardization] program]]]

which violates the maximum right depth constraint. He was unable, he says, to discover any other counter-examples.

There are, however, many more counter examples. Figure 8 lists several along with the structures which seem appropriate to me. What I find interesting is the fact that a very high proportion of long sequences of nouns do observe this constraint. This is a fact which I would like to explain. It might be possible to characterize the kinds of relationships for which it is more likely to find a violation of the constraint. It may be that the more 'primitive relationships are more amenable to structures which do not obey the constraint (e.g. relationships like source, time or location). It is also likely that the constraint is a side-effect of the processing strategy which is used by people in attempting to interpret long sequences of nouns. If this is true, then it would be an important heuristic to capture in any computer program which attempts to do the same, if only to rank the potential interpretations with respect to the probability of matching the speakers intended meaning.

number of nouns	$f(n)$ number of unconstrained trees	$f'(n)$ number of constrained trees
1	0	1
2	1	1
3	2	2
4	5	4
5	14	8
6	42	16
7	132	32
8	429	64
9	1430	128
10	4862	256

The number of constrained vs. unconstrained trees

figure 7

[Aluminum [automobile [water pumps]]]
["J.C. Whitney" [Automobile [water pumps]]]
[January [[automobile [[water pump] cover]] shipments]]
[[solid state] [RCA [color television]]]
[[back yard] [brick [dog house]]]
[Illinois [ERA [voice vote]]]
[prototype [MIT [LISP machine]]]
[plastic [toy [fire truck]]]

Examples which violate the Marcus Constraint

figure 8

III

The ApproachIII.1 Rule Representation

The basic mechanism of semantic interpretation will be driven by interpretation rules. These rules will be represented by frames and organized into the abstraction hierarchy. The use of frames to represent the interpretation rules will have several benefits.

First, this facilitates experimentation with the kinds of knowledge that will go into a rule. The addition or deletion of attributes of rules can be done simply by adding or removing slots from the generic rule frame. Moreover, the information stored in the rule frames can be easily augmented with ancillary information, such as the contexts in which the information is important.

Second, this representation allows the system to treat the rules as formal objects which can be the object of inference and manipulation. This facilitates the writing of meta-level rules, such as the rules for sets described in an earlier section of this proposal.

Finally, organizing the rules into an abstraction hierarchy aids the recognition of regularities. It also provides one way to restrict the application of a rule if a more specific rule is found to apply.

III.2 Classes of Rules

I anticipate using several general classes of interpretation rules for the interpretation of noun-noun modification. Although there is some overlap between these classes, I believe it is fruitful to think of them independently.

The first class I refer to as idiomatic rules. These rules will typically match surface lexical items directly. For example, the navy refers to a plane which has a very poor maintenance record as a "hanger queen". A rule to interpret this phrase would have a pattern which require an exact match to the words "hanger" and "queen".

The second class consists of productive rules. These rules attempt to capture forms of modification which are productive in the sense of defining a general pattern which can produce many instantiations. An example from this set would be a rule which attempted to view the modified noun as an artifact

and the modifying noun as a raw-material and produced the interpretation in which the underlying relationship was something like made of.

A third class of rules is based on procedures which analyze the representations of the concepts for the modifier and modified noun and attempt to discover an appropriate relationship between them. Many of these rules will be useful for analyzing compounds which contained nominalized verbs. One of their primary sources of knowledge will be the case frame associated with the verb. As an example, suppose we wish to interpret the phrases "delivery man" and "delivery truck". In each case, we note the "delivery" refers to the "deliver" event. The case frame for this event will have roles associated with it such as agent, object, instrument, time, etc. Associated with each case are specifications for the requirements and preferences for the role fillers. In interpreting the first phrase we would discover that there is a good match between "man" and the preference for the agent role (which might state that it prefers to be filled with a concept matching a person or an organization (e.g. UPS). In the interpretation of the second phrase, we would discover that "truck" matches the preference of the instrument role.

A final class of rules will be used to handle the difficult case of modification in anaphora. When noun-noun modification is being used in a definite description used anaphorically, the relationship between the modified noun and the modifier can be almost anything. What is required is a search of the discourse context for referents which involve a relationship between the two nouns.

IV

Scope of the proposed work

The work will be divided into four sections: a discussion of the theoretical issues, an exploration of implementational issues, the construction and testing of an experimental system, and a discussion of the impact of this work on several related areas.

IV.1 Theoretical Issues and Problems.

The following list gives some of the theoretical issues and problems that I hope to address.

(1) Underlying semantic representation

Any project involving semantic interpretation should be founded upon a representation which is logically adequate. Although this seems an obvious requirement, people are still building systems around logically incomplete formalisms. A criterion which is at least as important is the expressiveness of the underlying representation. The representation should make it easy to encode various kinds of knowledge, especially knowledge cast in a procedural framework. In addition, I believe it should support multiple representations or viewpoints of the same concept.

An important result of this work will be a better understanding of the kinds of knowledge that must be represented in order to handle the difficult problem of noun-noun modification with any degree of closure. Once the kinds of knowledge are identified, the issue of how it is represented is clearer.

(2) Enumerating typical forms of noun-noun modification

Many examples of simple noun-noun modification can be covered by a small class of productive rules. For example, most words referring to physical objects can be modified by words which describing the raw material or components out of which the object is constructed. Thus a single rule can be used in the interpretation of such phrases as:

rubber ball
vodka martini
concrete boat
leather coat

Other examples of very productive modification rule classes cover such relations as PART-OF (e.g. engine housing), LOCATION (e.g. urban riots, country roads) and TIME-OF (e.g. the January meeting).

(3) Selecting the proper sense of polysemous words

The general problem of word polysemy has been studied for verbs and their objects [FININ76]. In the context of noun modification it is just as problematical. In general, the sense of the modified word can depend on the potential senses of its modifiers, as in:

twelve inch wrench	wrench as a tool
sudden wrench	wrench as a action

and the sense of the modifying word can be selected by the potential senses of the head noun, as in:

nut driver	nut as a bolt fastener
nut shell	nut as an edible seed

Given an appropriate discourse context, the evidence at such a local level can be, of course, overridden. For example, D. L. Waltz has proposed to me the following sentence "All the nut drivers are on strike at the state mental hospital".

(4) recognizing and understanding idioms

I view idioms as being at one end of a spectrum of linguistic patterns which range from fixed, word-for-word idioms to complete interpretation by analysis of constituents. Thus, in my system, idioms are easily recognized by rules which match the surface lexical items.

(5) Mundane vs. novel language use

In my view, the great majority of our language consists of canned patterns and phrases. Collectively and individually we develop patterns of speech which become habitual [BECKER]. I refer to this as mundane language use.

We repeatedly use that same words and phrases to describe common concepts and events. A priori, there is no logical reason for this to be so. If our language facility were purely analytic/synthetic, we would expect to observe a wide range of descriptions for a single concept. Instead, we find that common concepts acquire stylized (or even hackneyed) descriptions. Often this description is promoted to the rank of a multi-word name (e.g. a compound noun) or an idiom.

Note that by mundane language use I do not mean common or un-specialized speech. In fact, I believe that this effect is even stronger in specialized language. To the layman, specialized language appears as an incomprehensible string of jargon [examples from CS]. In this specialized speech one finds many new words, idioms and phrases. To compound the problem, common words are given new meanings or have particular components of their standard interpretation emphasized.

The prevalence of mundane language has benefits, of course, such as a potential saving for processing time and a reduced danger of misinterpretation.

(6) Implications for learning the meaning of new words and phrases

Much of the linguistic research on nominal compounding has come from examining the process of coining novel noun compounds and creative nominalization. I hope that the more general interpretation rules I develop might point the way for the understanding of truly novel concept modification.

(7) Understanding analogy and metaphor

The problem of understanding analogy and metaphor has points in common to that of understanding new words and phrases. In both cases the normal interpretive mechanisms fail to result in an appropriate interpretation. I believe the basic processes which are hypothesized to be used in understanding metaphor and analogy may also be used to interpret noun-noun modification. This is an idea I would like to explore if time permits.

(8) sensitivity to context: linguistic, textual and pragmatic

An important factor in any semantic interpretation system is how the effects of context can be integrated with the semantic interpretation rules. The notion of context can be thought of at many different levels. In interpreting an instance of noun modification we are at a very local level and must be sensitive to the context of the enclosing sentence and the overall discourse context. One must also be sensitive to what I call a pragmatic context. By this I mean the set of pragmatic facts relevant to the goals and intentions of the user as well as the known limitations and capabilities of the computer system.

IV.2 Implementational Issues

A major result of this research will be a computer program which will interpret basic noun modification in noun phrases. The program will be one component of the JETS question answering system now under development. This program will act as a semantic specialist whose domain is the interpretation of certain kinds of basic noun modification.

The program will be given a parsed noun phrase and will build a semantic representation for it. If more than one interpretation is appropriate, a list of candidate interpretations will be generated. The list will be ordered by heuristics which measure the probability that a candidate is the intended interpretation.

The details of what an adequate meaning representation might be is an ongoing concern of our JETS research group. Over the next few months we hope to arrive at a preliminary design. One of the early decisions is that the meaning representation refer to the events and objects described by the data base rather than the information in the data base itself. In other words, our meaning representation will be much broader than the actual data base information requires. Our goal is to enable the translation from the

underlying meaning representation into a formal query by a simple mechanistic program.

Some of the implementational issues I foresee are:

(1) The design of a knowledge representation system

The FRL representation system merely provides a useful set of primitive functions for creating and manipulating a certain type of general data structure. It specifically does not provide a theory of representation or even representational semantics. My work will require me to give some thought to the general issues of what a general representational system should and should not do and then to implement my own variety on top of FRL.

(2) Efficient retrieval of rules

(3) Efficient representation of contexts

IV.3 Peripheral Issues

There are many related issues that my work will not address but may have an impact on. If time permits I would like to briefly explore some of these areas. Some candidates that come to mind are:

(1) The interaction with a syntactic parser

(2) the interaction with a Database specialist.

(3) Implications for the analysis of other forms of modification

V

References

- [BOBROW] Bobrow, D.G. and Winograd, T.W.; An Overview of KRL-0, a Knowledge Representation Language; Proceedings of the 5th International Joint Conference on Artificial Intelligence; August 1977;
- [BORGIDA] Borgida, A. T.; Topics in the Understanding of English Sentences by Computer; TR78, Department of Computer Science, University of Toronto; February 1975;
- [BRACHMAN] Brachman, R. J.; What's in a Concept: Structural Foundations for Semantic Networks; International Journal of Man-Machine Studies, vol 9., pp. 127-152; 1977;
- [BECKER] Becker, Joseph D.; The Phrasal Lexicon; Theoretical Issues in Natural language Processing - 1; June 1975.;
- [BULLWINK] Bullwinkle, C.; Levels of Complexity in Discourse for Anaphora Disambiguation and Speech Act Interpretation; Proceedings of the 5th International Joint Conference on Artificial Intelligence; August 1977;
- [CHANG] Chang, C. L.; Finding Missing Joins for Incomplete Queries in Relational Data Bases; Report RJ2145, IBM Research Laboratory, San Jose; February 1978;
- [CODD] Codd, E. F., Arnold, R.S., Cadiou, J-M., Chang, C. L. and Roussopoulos, N.; RENDEZVOUS Version 1: An Experimental English-Language Query Formulation System for Casual Users of Relational Data Bases; Report RJ2144, IBM Research Laboratory, San Jose; January 1978;
- [FILLMO68] Fillmore, Charles J.; The Case for Case; in Bach and Harms (eds), Universals in Linguistic Theory, Holt, Reinhart and Winston, New York; 1968;
- [FILLMO77] Fillmore, Charles J.; The Case for Case Reopened; in Syntax and Semantics; 1977;
- [FININ76] Finin, Timothy W.; Determining Verb Sense; unpublished working paper; 1976;
- [FININ79] Finin, Timothy, Goodman, Bradley and Tennant, Harry; JETS: Achieving Completeness through Coverage and Closure; Proceedings of the 6th International Joint Conference on Artificial Intelligence; August 1979 (to appear);

- [GERSHMAN] Gershman, A.V.; Conceptual Analysis of Noun Groups in English; Proceedings of the 5th International Joint Conference on Artificial Intelligence; August, 1977;
- [GLEITMAN] Gleitman, Lila R. and Gleitman, Henry; Phrase and Paraphrase; W. Norton, New York; 1970;
- [LEES68] Lees, Robert B.; The Grammar of English Nominalizations; NOSOURCE; 1968;
- [LEES70] Lees, Robert B; Some Problems in the Gramatical Analysis of English Nominal Compounds; in Progress in Linguistics, Mouton; 1970;
- [LEVI] Levi, Judith N.; The Syntax and Semantics of Complex Nominals; in press;
- [MARCUS] Marcus, Mitchell; A theory of Syntactic Recognition for Natural Language; unpublished Ph.D. thesis, MIT; 1977;
- [MCDONALD] McDonald, D. and Hayes-Roth, F.; Inferential Searches of Knowledge Networks as an Approach to Extensible Language Understanding Systems; in Waterman and Hayes-Roth (eds.) Pattern-Directed Inference Systems, Academic Press; 1978;
- [MINSKY] Minsky, M.; A Framework for Representation Knowledge; in Winston (ed.) The Psychology of Computer Vision, McGraw Hill, New York; 1975;
- [RANDOM] The Random House Dictionary of the English Language, College Edition; Random House; 1968;
- [NALDA] NALDA (Naval Air Logistics Data Analysis) System Data Requirements Determination Report; Naval Aviation Integrated Support Center, Patuxent River, Maryland; 1975;
- [RHYNE] Rhyne, J. R.; A Lexical Process Model of Nominal Compounding in English; American Journal of Computational Linguistics, microfiche 33; NODATE;
- [ROBERTS] Roberts, B. R., and Goldstein, I. P.; The FRL Manual; MIT-AI Memo 409, Massachusetts Institute of Technology Artificial Intelligence Laboratory; September 1977;
- [SEARLE] Searle, J.R.; Metaphore; in press;
- [SIMMONS] Simmons, R.; On managing Sentence Meaning; CAI Laboratory report NL-6, Computer Science Department, The University of Texas, Austin; 1972;
- [SOWA] Sowa, J. F.; Conceptual Graphs for a Data Base Interface; IBM Jour. of Research Developement; July 1975;

- [TENNAN78] Tennant, Harry; The PLANES Database; Working paper 14, Advanced Automation Group, Coordinated Science Laboratory, University of Illinois; June 1978;
- [TENNAN79] Tennant, Harry; Experience with the Evaluation of Natural Language Question Answerers; Proceedings of the 6th International Joint Conference on Artificial Intelligence; August 1979 (to appear);
- [WALTZ76] Waltz, D.L., Finin, T.W., Green, F., Conrad, F., Goodman, B. and Hadden, G.; The PLANES system: Natural Language Access to a Large Data Base; Technical Report T-34, Coordinated Science Laboratory, University of Illinois; November 1976;
- [WALTZ77] Waltz, D. L. and B. A. Goodman; Writing a Natural Language Data Base System; Proceedings of the 5th International Joint Conference on Artificial Intelligence; 1977;
- [WALTZ78] Waltz, D. L.; An English Language Question Answering System for a Large Relational Data Base; CACM, vol. 21, pp. 526-539.; July, 1978;
- [WALTZ79] Waltz, D. L.; Relating Images, Concepts and Words; Proceedings of the 6th International Joint Conference on Artificial Intelligence; August 1979 (to appear);
- [WOODS72] Woods, W.A., Kaplan, R.M., Nash-Webber, B.; The Lunar Sciences Natural Language Information System: Final report; Bolt Beranek and Newman report No. 2378; June 1972;
- [WOODS75] Woods, W.A.; Whats in a Link: Foundations for Semantic Networks; in D.G. Bobrow and A.M. Collins, eds, Representation and Understanding: Studies in Cognitive Science, Academic Press; 1975;
- [WOODS77] Woods, W. A.; A Personal View of Natural Language Understanding; SIGART Newsletter, number 61, pp. 17-20; February 1977;