

AD-A138 455

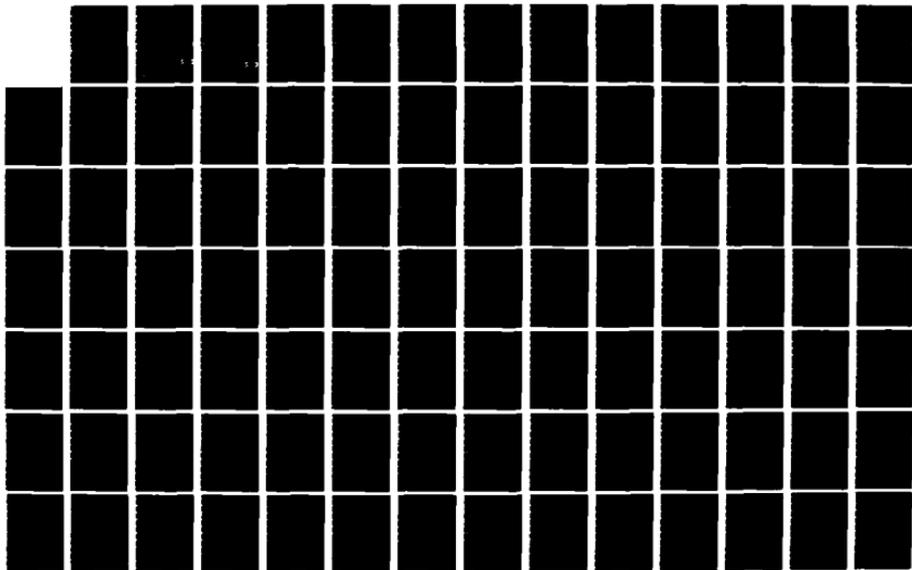
NATURAL LANGUAGE PROCESSOR AS A UNIVERSAL FRONT END TO
EXPERT SYSTEMS(U) AIR FORCE INST OF TECH
WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING
D L BROOKE DEC 83 AFIT/GE/EE/83D-13

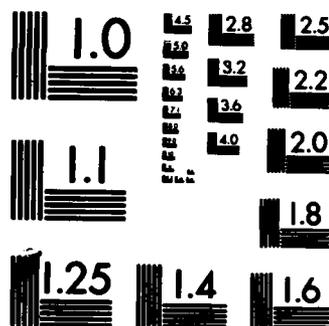
1/2

UNCLASSIFIED

F/G 9/2

NL





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

AD A138455



①

Natural Language Processor
as a Universal
Front End To Expert Systems

Thesis

Deborah L. Brooke
Capt USAF

AFIT/GE/EE/83D-13

DISTRIBUTION STATEMENT A

Approved for public release
Distribution Unlimited

DTIC
ELECTE
MAR 1 1984

S

B

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

DTIC FILE COPY

Wright-Patterson Air Force Base, Ohio

84 02 29 04

**Natural Language Processor
as a Universal
Front End To Expert Systems**

Thesis

**Deborah L. Brooke
Capt USAF**

AFIT/GE/EE/83D-13

**DTIC
ELECTE
S MAR 1 1984 D
B**

Approved for public release; distribution unlimited

EE
AFIT/GE/~~ENG~~/83D-13

**NATURAL LANGUAGE PROCESSOR AS A UNIVERSAL
FRONT END TO EXPERT SYSTEMS**

THESIS

**Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology**

Air University

**In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Electrical Engineering**

Deborah L. Brooke, B.S.E.E.

Captain, USAF

December 1983

Approved for public release; distribution unlimited

Preface

The purpose of this study is to specify a Natural Language Processor which is to be used as a front-end to expert systems used by the United States Air Force. It is important to specify a standard, modular interface early in the development of expert system/natural language interfaces to ensure compatibility from system to system throughout the Air Force. I hope this work proves useful to that end.

I wish to thank my thesis advisor, Capt Rob Milne, for his patience and assistance throughout the writing of this thesis. I sincerely appreciate the encouragement of Major Ken Castor and Dr John Muller whose confidence in me enabled me to complete my work. I also wish to thank Capt Gregg Gunsch and his wife Cherry for their support and their help with my daughter when my work got tough.



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
<i>PER CALL JC</i>	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
<i>A-1</i>	

Table of Contents

	Page
Preface.....	ii
Abstract.....	iv
I. Introduction.....	I - 1
Background.....	I - 6
Conclusions.....	I - 8
II. Air Force Requirements.....	II- 1
III. Components of Natural Language.....	III- 1
User.....	III- 2
Character of the Dialog.....	III- 3
Utterance Structure.....	III- 4
Parsing.....	III- 7
Problematic Use of Natural Language	III- 8
IV. Introduction to Expert Systems.....	IV- 1
SCHOLAR.....	IV- 2
INTELLECT.....	IV- 4
KNOBS.....	IV- 7
ISOBAR.....	IV- 9
CRYSALIS.....	IV-10
PROSPECTOR.....	IV-11
ESCA SPECTRA INTERPRETER.....	IV-12
Requirements on Expert Systems.....	IV-14
V. The Natural Language Processor - NLP...	V - 1
Input Facilities.....	V - 1
Output - Text Generation.....	V - 5
Interfacing NLP With Expert System.	V - 6
VI. Summary.....	VI- 1
Appendix A: Dialogs With Expert Systems....	A - 1
Appendix B: Air Force Vocabulary.....	B - 1
Bibliography.....	BIB- 1

Abstract

This thesis specifies a list of features which must be included in a transportable natural language processor (NLP) used as a front-end to expert systems used by Air Force personnel. It also places requirements on the expert systems themselves in order to interface with such a NLP. The state-of-the-art in natural language processing is such that it can support such an NLP/expert system interface.

The following features must be provided by a NLP for expert systems which may be used by the Air Force.

1. Process one and two sentence English utterances from a user.
2. Process utterances that most people would understand if they read them in context. This includes well and ill-formed utterances.
3. Resolve pronominal references, ellipsis, and anaphora.
4. Error handling facility to resolve lexical ambiguity of input.
5. Generate output text for its error handling facility, the expert system domain answers, and the clarification facility of the expert system.
6. Manage a lexicon with words and phrases that most users need in the given expert system domain.

7. Allow user to upgrade the lexicon.
8. Spelling correction.
9. Reasonable processing times for each input.
10. Output a fixed, logical format to be used as input to all expert systems.

The following features should be included in any expert system which is to be interfaced with an NLP for Air Force expert systems:

1. Knowledge base organization which can accept a fixed, logical format from a modular NLP and be resolved using the rules of predicate calculus.
2. Knowledge base which "knows" facts about its knowledge domain as well as what it doesn't know about its knowledge domain.
3. Knowledge about what is not included as part of its knowledge domain.
4. Access rules which support the NLP to communicate this knowledge.
5. Detection of logical inconsistencies and ambiguities.
6. Question/Answer facility for clarifications.
7. Reasonable processing times for each input and output.

INTRODUCTION

1.0 Problem

To accomplish its mission the Air Force has many people working on many specialized tasks which are all integrated into the Air Force defense mission. Many of the people who perform these specialized tasks are classified as "experts" in their fields. One of the ubiquitous tools used by personnel to accomplish their individual missions is the computer.

Computers have become fully integrated into many different Air Force activities. One class of specialized computer programs called, expert systems, shows promise of taking computer use beyond number crunching and data storage. An expert system performs the function of a personal, on-line expert for the computer user. It is possible to have as many different kinds of "expert systems" as there are topics to be an "expert" on. Many expert systems have already been written and several of them will be discussed in this thesis. One of the problems with expert systems is that they have narrow topic orientations, i.e. a very specialized knowledge domain. This makes them each unique in their usage. Quite often they have an input language designed specifically for the individual expert system. For one person to use

many expert systems, they may have to learn many different computer languages. If a person could learn all these languages, they would have a tremendous amount of knowledge at their disposal. The topic of this thesis is not expert systems, however.

It is important to be able to use a computer with as little training as possible to conserve personnel resources. Many computer systems, are simple enough to use without lengthy training, for example, word processors. Expert systems, however, can be very complicated programs and have narrow domains. A means of accessing these programs without lengthy training is to allow access to the systems via natural language, which for the United States Air Force, is English.

Using English as a computer input language allows people of all fields access to powerful expert systems. It puts the responsibility of knowing how to extract information on the computer instead of on the user. Any system which used natural language as the input would have many users, not just a select few.

1.1 Scope

This thesis is the specification of a transportable natural language processor interface (NLP) to expert

systems. The thesis will:

1. Show how an NLP should interface with expert systems,
2. Show what expert systems should do to use the NLP.

There are two facets of this task which must be defined up front: the term "transportable" and the acceptable degree of natural language processing. Transportability implies that you can pick up the NLP and use it wherever you need to. The degree to which the nlp is transportable must be defined. The other problem is that a computer natural language processor does not necessarily process natural language the same way, or to the same degree, that humans do. It is necessary to define what is meant when the NLP "understands" the "meaning" of an utterance.

1.1.1 Transportability

The NLP uses the concept of transportability to save work. When an expert system uses natural language as its input, the rules of the language do not change even if the user or the knowledge domain changes. Most of the elements of natural language processing are common from application to application. Therefore, one NLP should be

able to process the input for all systems which use natural language as input. A modular design where the natural language processor forms one part of the overall system, and the expert system the other, will save much design work. It will also allow users to become familiar with only one natural language processor instead of several. The only difference from one system to the next will be the expert's domain.

The NLP will need to be adapted to each expert system. The NLP and the expert system, to which it is interfaced, must share the same interpretations of words. For example, when a user asks a question about ships, both systems must know whether the user means the verb "to ship", or the noun "ship". To date, most natural language processors were designed with a specific expert system in mind, and the expert system and its interface with the user were not separable. In a modular design, the two ends of the system are designed separately, and must be made to work together after they are each built. The degree of transportability of the NLP will vary depending on how much modification work goes into adapting the NLP to its "host" expert system.

The purpose of this specification is to simplify the task of making the expert system and NLP work together by giving the designers of each information about what to

expect from the other end of the system. Depending on the knowledge domain of the expert system, the NLP could be adapted by either a NLP expert or the domain expert. This system specification ASSUMES that a NLP expert will do the adaptation task. It is difficult enough, given the "state of the art", to design a practical system which "understands" domain commands. It is still more difficult to design a NLP which "understands" commands about what you want it to "understand"; and the task of understanding the domain still has to be done. The TEAM NLP system of SRI is being designed to be modified by the domain expert.[Grosz, 1982] The ROBOT NLP system by Harris assumes the NLP expert does the modification. ROBOT is currently commercially available. TEAM is still under development.

1.1.2 Understanding and Meaning

The NLP will be an input processor to a language which is not understood by those who study it. No one has been able to write a specification of "natural language" which satisfactorily embodies the essence of what natural language is and how it is used. Experts in linguistics, philosophy, psychology, and artificial intelligence all define "natural language" from different perspectives. Concepts such as "understanding" and "meaning" have no universally accepted definitions, and

these concepts form part of the foundation of natural language usage. Do to a lack of understanding, instead of defining these words, the best we can do is work in terms of levels of understanding and meaning. In this thesis then, an intuitive feeling of what "understanding" is will be relied on and the level of understanding will be defined.

For the purposes of the NLP, the level of "understanding" is defined as: assigning the correct "meaning" to an utterance and demonstrated by the NLP producing the output desired by the user. The NLP is limited to assigning meaning to utterances only one or two sentences long. As will be seen, Chapter 3 and Appendix I, most expert systems do not require more than that. The NLP is not at all capable of assigning meaning to paragraphs of text.

1.2 Background

As stated before, a computer expert system performs the function of a human expert. Expert systems have been written to aid in learning, ICAI systems, to act as consultants, MYCIN and PROSPECTOR, to perform analyses, CRYSLIS, and to function as a general assistant, data-base access with systems such as ROBOT. Systems are envisioned which will function as intellegent robots which

perform hazardous duties, and as strategic planners.[Brown, et al, 1982] Brown, et al list and analyse many expert system ideas with direct military application.

In general, the expert system is composed of four parts:

1. Input section,
2. Output section,
3. Knowledge base,
4. Access rules to knowledge base.

The first section of the expert system may take as input: requests, commands, declarations of fact, or a combination of these depending on the nature of the knowledge domain.

The output section of the expert system generates answers to a user's input. This output is used as input to a text generation facility of the NLP for display to the user.

The knowledge base of the expert system is the collection of definitions and productions which form the expert's domain. Productions are rules used by the expert system to produce an answer, for example, the rule that "1 plus 1 equal 2".

The access rules of the expert system match a user's input to the information the system contained in its knowledge base.

The NLP works directly with the first two sections of the expert system. The organization of the latter two sections of the expert system must support this type of arrangement.

It is necessary to specify what an expert system should do in order to use natural language as the input language BEFORE too many expert systems are fielded. This way all systems can be built to the same standard. The number of expert system under development is growing. It is necessary to set these standards while there is time to adhere to them. Each of the parts of the expert system must be designed with the idea of a NLP as front-end in mind.

1.3 Conclusions

There are two types of knowledge embodied in a NLP/expert system interface: linguistic knowledge and concept knowledge. Concept knowledge contains both domain knowledge as well as the knowledge which communicates it. The NLP is responsible for communicating and managing the

linguistic knowledge. The expert system is responsible for communicating and managing the conceptual knowledge. The NLP and the expert systems to which it attaches must meet the following criteria which will be detailed in the rest of the thesis:

1.3.1 Natural Language Processor

1. Process one and two sentence utterances from a user,
2. Process well-formed AND ill-formed English utterances,
3. Resolve pronoun references, ellipsis and anaphora,
4. Generate output text,
5. Manage a lexicon with words and phrases which accepts most user inputs,
6. Allow user to upgrade the lexicon,
7. Spelling correction,
8. Error handling facility when it can't resolve the input at all,
9. Reasonable processing times for each input.

1.3.2 Expert System

1. Knowledge base organization to support the above,
2. Knowledge about its data-base and what it doesn't

know about it,

3. Knowledge about the limits of its domain and what is not included in it,

4. Access rules which support the NLP to communicate this knowledge,

5. Detection of logical inconsistencies and ambiguities,

6. Question/Answer facility for clarifications,

7. Reasonable processing times for each input and output.

1.4 Organization of Thesis

Chapter 2 analyses the Air Force's requirements for expert systems and natural language processing.

Chapter 3 describes the components of natural language dialogue. This includes the user, and the features of natural language usage such as grammar, syntax, semantics, and vocabulary. The problems posed to natural language processing are discussed.

Chapter 4 describes the expert systems with which this processor interfaces and the constraints placed on these systems so that they may interface with a modular natural language processor. Dialogues with the expert systems are provided (see Appendix I) and examined for

types of input and output to be handled as well as the vocabulary used. Requirements that the expert system must fulfill in order to interface to an expert system are detailed.

Chapter 5 describes the natural language processor itself and the features it must have. The kinds of language phenomenon it must handle are discussed, the vocabulary, the grammar, text generation, and connection to the expert system are detailed.

Chapter 6 contains the conclusion, recommendations for future study, and other possible problems which are not taken into consideration in this study.

CHAPTER 2

2.0 Air Force Requirements

This chapter examines some of the characteristics of Air Force tasks, and delineates some of the requirements generated by these characteristics. It then shows how expert systems and natural language processing satisfy some of these requirements. In particular, vocabulary which must be included in a NLP for the Air Force is discussed. The list of specific words is in Appendix B.

2.1 Characteristics

In general, the Air Force is a large organization of military persons. Without going into Air Force doctrine and specific mission, suffice it to say that the majority of tasks in the Air Force are technical or clerical. Each person in the Air Force has a special field in which they work. Most of the personnel require training of some sort for their specialty. In order to minimize these training requirements, the tasks and the equipment that personnel work with, in principle, have standard features so that once a person is trained to operate with a type of equipment or perform a certain task, they do not have to be retrained to account for individual differences in the operation.

When equipment is procured, it must operate with many other pieces of equipment already in the field. Each piece of equipment procured has standard, interchangeable features. The equipment is usually bought in very large numbers to be used by very large numbers of Air Force persons. Each piece of equipment will probably be operated by many different persons. No one person is issued a piece of equipment that they will keep for their very own other than on a short term basis.

2.2 Requirements

As seen from the above description, the Air Force consists of many standard and interchangeable parts which fit together in a modular fashion. This characteristic means that anything added to the Air Force "system" must be compatible with what is already there and be able to accommodate what will come afterwards. This is the reason that an expert system or NLP added to the Air Force inventory must be modular and standard. This is also the reason for the need for a standard Air Force vocabulary.

The Air Force standard vocabulary was found by examining representative dialogs and writings which occur in the Air Force. These writings included: Air Force Standard Dictionary of Terms which listed all the "officially" recognized Air Force words. Four issues of

Flying and Airman magazine were examined to see what words made up Air Force "lingo". Most of the Air Force words were picked out after the first two magazines and the other magazines merely used the same words over again. "United States Military Posture For FY 1983" was also examined closely for more Air Force terms. Several miscellaneous books selected randomly from library shelves were also examined, but were not used because they appeared to repeat the same words over again. A fairly complete list of Air Force terms appears in Appendix B of this thesis.

Not all of the Air Force "terms" which appeared in these documents was included in the list in Appendix B. Some of these terms, such as "aircraft arresting hook cable tie-down", were so "domain" specific that they found very limited use and should be added in at a later time should a user find it necessary. However, the list in Appendix B should contain every word, and unfortunately more, that an Air Force expert system user should ever require.

CHAPTER 3

3.0 Components of Natural Language

The purpose of a Natural Language Processor is to allow a user to express herself in the "manner to which she is accustomed".[Tennant 1981:50] In order to specify a NLP, the nature of natural language must be specified. It is this nature that must be accommodated.

A complete specification of natural language dialog has never been made to the satisfaction of the phsychologists, philosophers, linguists, and AI researchers. What follows is not complete; it is a general description of the major attributes of natural language usage. A much more complete "Taxonomy" can be found in [Tennant 1981:82-84, 93-100]

Any communication can be viewed as having three major components: a sender, a receiver, and the message. A human/computer natural language dialogue can be viewed the same way with three analogous components: a user, the NLP/expert system, and the natural language input. This chapter describes the characteristics of such a human/computer dialogue. First the character of the user is described. Then the character of natural language input will be described which can be broken down into

vocabulary, syntax, semantics, and grammar; each of which will be defined. The input can also be described as: utterances, sentences, and entire dialogues composed of paragraphs and a central idea. We will assume the input is in the form of "utterances" whether or not the utterance comes in the form of one or more "sentences". Last, this chapter describes methods used to process the natural language input. The user types it into a keyboard, a parser takes the utterance apart, "formats" it, and passes it to the expert system.

3.1 The User

For the purposes of the NLP design, it is assumed that the user is cooperative and has nothing to gain by trying to lie or cheat the NLP or the expert system. The user's level of sophistication varies from naive to an expert in the expert system's knowledge domain. The user is assumed to be associated with the military and is not necessarily at all interested in computers themselves or the details ordinary computers require for interaction.

The user's purpose when using an expert system is defined by its knowledge domain, therefore, the ensuing dialog is goal-directed. The expert system always plays the same role. The function of the discourse and the domain of the discourse will be the same for all users of

a given expert system. The user of an expert system will assume that utterance "interpretation is literal whereby only propositional content is determined." [Linde 81] This assumption about the character and purpose of the user eliminates the requirement to deal with such natural language phenomena as metaphor and much "world knowledge" information. A user of the PROSPECTOR system will not expect PROSPECTOR to know that people ride "on a bus" and "in a car". The user of an expert system is not looking for idle chit chat or to practice her rhetorical devices.

The user can be idiosyncratic and justifiably expects the NLP to accommodate her.[Tennant 1981:45] He describes a situation where the user wanted to use the phrase to "deal with". The natural language processing system he was using, Automatic Advisor, was not familiar with the term so the user tennaciously tried to define "deal with" to the system to force it to adapt to him; not the other way around.

3.2 Character of the Dialog

3.2.1 Length Of Discourse

The NLP will be required to process one or two sentence utterances. In Chapter 3 it will be shown that a

dialogue in which an idea is pursued in extended text, such as paragraphs, is not necessary for expert systems. The ideas which surround such a dialog are very complex and involve concepts of "meaning" and "perception". [Schank 1980] [Lehnert 1980] [Waltz 1979]

The NLP only processes dialogue with expert systems. This restriction narrows the range of inputs from "anything a person can dream up" to the restricted knowledge domain of an expert system. A user of an expert system such as Shrdlu [Winston 1979] only expected to move colored blocks around a table. The knowledge domain of the expert system is reflected into the user's expectations and, therefore, into her dialogue.

3.2.2 "Utterance" Structure

Utterances are characterized by their vocabulary, syntax, semantics, and grammar. They can be ill-formed or well-formed sentences. This section defines vocabulary, syntax, semantics, and grammar as well as some methods of deviant linguistic usage.

3.2.2.1 Vocabulary

English words can be categorized as closed-class or as open-class words. Open-class words are those used no

matter what the topic under discussion. These are the words like: and, the, will, no, yes, go, have, and if. "Open-class words are the common noun, verb, or adjective". [Milne 1983] They are associated with a specific knowledge domain and have meanings which change with the domain. Words such as: injection, decibel, and noise, mean something different in the "electronics domain" than they do at a cocktail party.

It is a straightforward process to determine what words are needed in the lexicon of an expert system. A test has to be performed over a reasonable period of time to see what words are used when working with an expert system. [Thomas 1977] The lexicon for the RUS parser contains an extensive list of the closed-class words used by persons during sessions with expert systems. [Bobrow 1979]

A list of the open-class words are obtained the same way. Many of the words can be obtained from the literature of the knowledge domain, but most come from observing persons using them during sessions with the expert system. Chapter 3 treats the open-class words, and a list of the words used in the USAF is in Appendix II.

3.2.2.2 Syntax

Syntax is the structure of the sentence or utterance. It classifies a word based on its position within a sentence. For example, English syntax requires that adjectives precede the noun they modify. Someone listening to a sentence could classify a word as a noun, verb, or adjective based on its location in the sentence.

3.2.2.3 Semantics

The semantics of a word are its meanings. The meaning of a sentence can vary from context to context. For example, "Mary saw it." "Saw" can "mean" any number of things depending on the context in which it was uttered. If the sentence were a command, someone may be asking Mary to "saw" a piece of wood. Or, the sentence may be stating that Mary "observed it". These interpretations of "saw" are the word's semantics.

Note that a sentence can be "syntactically" correct, but still have no meaning. A famous example is the sentence: "Colorless green ideas sleep furiously." It is syntactically perfect but a semantic nightmare because it has no meaning.

3.2.2.4 Grammar

Grammar is the set of rules which govern how words may be used together. It relates the syntax and semantics of a word together. In English we make rules like: all sentences have to have a noun for a subject and a verb for a predicate, and adjectives sit on the left hand side of the noun they modify. There are other grammars besides English such as simpler "syntactic" or "semantic" grammars. A syntactic grammar assigns words meaning based strictly on the location in the utterance. A semantic grammar may not have nouns or verbs at all and a word will have the same "meaning" whenever it is used.

3.2.3 Parsing

Parsing is a process which converts the format of an input sentence into the format needed by the natural language processor. The rules that the parser uses to convert the input sentence into the proper output form are called the parser's grammar. A parser maps the words which are input by the user into a logical form that can be used by the expert system. The expert system will use the words provided by the NLP to provide the user with her request.[Rowley 1982]

An elementary method of parsing is simple "pattern

matching". The ELIZA system written by Weizenbaum in 1965 used pattern matching to process a user's natural language input. All a pattern matcher does is take a string such as <my sister has a cold> and match it with <my * has a *> where the "*" matches "sister" and "cold" in this pattern. (It is like wildcards in poker.) [Winston 1979] Other methods are presented in [Shank 1978] and [Hayes 1981].

Parsing would not be nearly as difficult as it is if people used English according to the rules of English. People use poor or wrong syntax, incomplete sentences, and often break off in mid-thought and try to pass it off as a whole thought. Since this is considered "natural", it must be dealt with in a practical way. Some of the problems encountered in parsing and theories to get around them are presented in [Birnbaum 1979] and [Levin 1979].

3.2.4 Problematic (Creative) Use Of Natural Language

"Natural" language is loose and full of ambiguities. This character has prevented philosophers, linguists, psychologist and AI researchers from forming a coherent, concise theory on how to interpret natural language. "There is no 'literal' meaning of a sentence; only the meaning a given listener places on it at the time of utterance." [Miller, 1976:704]

The NLP cannot assume that utterances are meaningful and well-formed. Attributes which have to be dealt with for the specification of a natural language processor are:

1. Pronominal references,
2. Ellipsis,
3. Anaphoric references,
4. Metaphor, simile, analogy,
5. Lexical peculiarities:
 - a. Sustitution,
 - b. Multiple meanings,
 - c. Mis-spelling.

3.2.4.1 Ellipsis

Ellipsis is a semantic phenomenon where a word or portion of a sentence or question is not explicitly stated but is referred back to in previous sentences, such as the following discourse:

"What are the specs on resistor R2?"
"What about R3?"
"C4?"

The subjects of the second and third sentences are "the specs of R3" and "the specs of C4". The practice of making the subject of your present sentence the same as

previous sentences without specifically stating so is known as ellipsis. The ellipsis can be resolved by matching the pattern of two adjacent sentences when the second sentence cannot be parsed. The pattern <What ... about R3?> matches the first sentence pattern <What R2?>. The parser fills in the words < ... are the specs on register....> for the second sentence so that it becomes, < What are the specs on resistor R3?>. The third sentence <...C4?> is matched to the pattern of <What about ...?> of the second sentence and then to the resolution of the second sentence to finally become; <What are the specs on C4?>.

Another form of ellipsis is the simple compound verb sentence: "Mary had a little lamb and drank a little wine." Mary is the subject of both verbs of this sentence. We have looked at one method, but there are alternate ways of resolving ellipsis. Some systems can resolve ellipsis but the method that is used results in anomalies and irregular coverage. For example, one researcher noticed that the expert system called PLANES could not resolve ellipsis when a conjunction was involved.[Tennent 1980:88] I am left wondering if the rules for handling conjunctions conflicts with rules for ellipsis resolution. Ellipsis frequently occurs in dialogues and shortens the number of words needed for discourse.

3.2.4.2 Pronominal Reference

Pronominal reference is merely the use of pronouns, with which all users of English are familiar. In a natural language processor it is a complicated event. The processor must figure out who, or what, the pronoun is referring to by going back to a previous sentence. The LIFER parser for the LADDER expert system, looks for an analogous pattern in the syntax of the user's last input, and then fills in the word that matches that pattern. [Hendrix 1977:10]

3.2.4.3 Anaphoric Reference

Anaphora is similar to ellipsis and pronominal reference in that it too refers back to a word or group of words already uttered. "The device of anaphora allows a speaker to identify an entity to a listener without giving a full description of the entity. The description given is not sufficient to distinguish the entity from everything that the listener knows about, but only to distinguish it from a much smaller group of entities directly related to the current state of the dialogue or discourse." [Hayes 1981:2] For example, "She dances better than he does." Instead of "She dances better than

he dances." Implementation of anaphoric resolution in a parser is similar to resolving pronouns and ellipsis. Methods of anaphoric resolution, and in particular her own theory, are explained in detail in [Webber 1978].

3.2.4.4 Metaphor, Simile, Analogy

Metaphor, simile, and analogy are all implementations of the phenomenon of describing a "thing" by relating its character to the character of another "thing". This phenomenon is very complicated and seems to have no end of variety in use. Expert systems are being experimented with to try to resolve these phenomena. A practical natural language processing system, at the present time, has no hope of handling them without driving the user bonkers with its requests for clarification and long processing times. For more information on techniques of resolution, see [Carbonell 1981].

3.2.4.5 Lexical Peculiarities

As if sentential abuse was not enough, words themselves are used in "unstandard" ways: substitution, multiple meaning, and misspelling. These phenomena can all be observed in the dialogs in Appendix I and are well documented throughout the literature.

3.2.4.5.1 Substitution

Substitution is simply using one word instead of another. This can cause problems in computer natural language processing because during lexical lookup, the processor must know that "Kennedy" also means "J F Kennedy". As documented by [Harris 1979] concerning the INTELLECT natural language processor, it can get even more complicated. For example, ME stands for the state of Maine. ME is also the first person singular objective case of "I". Harris apparently still doesn't know what to do about ME.

3.2.4.5.2 Multiple Meaning

Multiple word meanings is a mapping of one word into several meanings. A given word may function as a verb, as in "I'm going fishing." And, it can function as an adjective, as in "I need my fishing rod." This particular example is not very complex, but suppose the sentence read, "I need my fishing rod to go fishing." It is still worse as pointed out by [Finin 1977] for a phrase such as "...engine damage reports...". The parser must figure out that the entity "engine damage reports" is one entity, not the three entities: engines, damages, and reports. Another view of "multiple meaning" is the use of words as

metaphor which requires a whole expert system by itself to process. An interesting presentation of "novel language" processing is in [DeJong 1982].

3.2.4.5.3 Misspelling

We all know how to misspell. The parser's task is to guess when you mean to use specific word or if you misspelled it. LIFER's parser assumes you spelled the word correctly until the word is not in the lexicon. It will then tell you what word it assumes you meant (with the spelling corrected; see Appendix I) and proceed on that basis. This feedback is necessary in case you really did mean to spell a word the way you did, but the word is simply not listed in the lexicon. Entering a word which is not in the lexicon is a whole different problem.

3.2.4.6 Vague and Ambiguous Word Usage

Words have meanings as they are singly expressed as well as within a sentence. Putting the words together to express an idea is an art and not a science. A NLP has to work in the context of art not science.

3.2.4.6.1 Users do not think through what they want before expressing it. [Ehrenreich 1980] [Finin 1977]

This means that they don't ask for what they want even

though they think that they have. It is very difficult for people to unambiguously specify what it is they want. A NLP has to cooperate with the user in specifying what is needed. [Ehrenreich 1980] [Thomas 1977] A good example is presented in the following human(U)/computer(**) dialogue:

U: Are there summaries for January?

** Yes.

U: Could I have the January summaries?

** Yes.

U: I would like the January summaries.

** I understand.

U: Where are the January summaries?

** I don't understand that.

U: Can you give me the January summaries?

** Yes, I already told you that.

U: Would you please give me the summaries for January?

[Waltz 1981:25]

3.2.4.6.2 Novel word usage is common. One form of novel word usage is "engine damage reports" [Finin 1977] where a noun in this case is used a modifier. The lexicon contains descriptions of "engine", "damage", and "reports" but the concept of "engine damage reports" is nowhere to be found. [Waltz 1981] The phrase then becomes part of the "open-class" word group. Another form is the addition of prefixes and suffixes which change the meaning and spelling of the word so it may not be in the lexicon as written.

3.2.4.6.3 Users make sloppy use of quantifiers. Studies have shown that most people do not use quantifiers in the strict sense that logicians do. An example is "All B are C" and "All B are A", many people conclude that "All A are C" when it is actually true that only "Some A are C". [Thomas 1976] When a user asks, "Can this task be done in 4 hours?", she means "...within 4 hours". A system must be prepared to interpret the dialog as such. [Ehrenreich 1980]

3.2.4.6.4 Users make novel use of conditionals. Again, a user is not likely to use conditionals as a logician. [Thomas 1977] A sample dialog is given:

** May any Item be backordered?

User: Yes, as long as its available.

User: No, our price reductions are normally a result of being overstocked or because of special purchases that we make.

The user will answer yes or no but often adds a conditional to it. "A system must also be prepared to deal with an answer such as "yes, if X." where a direct computer test of X is impossible." [Thomas 1977: 26, 27]

Studies also show that the misuse of conditionals is reduced in the context of a dialog, although not eliminated. [Miller 1974] A non-programmer is likely to say "Put the red things in the box." as opposed to "If it is red, put it in the box".

3.2.4.6.5 Not all utterances are queries or commands.

Some are used to focus the dialog, for example, "I am now interested in A-7's." [Finin 1977] This is to be considered useful information.

3.3 Summary

This chapter discussed many of the linguistic phenomena that users are accustomed to. Not all of them have to be accounted for in a NLP. Chapter 4 will discuss which ones must be accommodated and which can be safely

ignored in implementation. Chapter 3 discusses the "conceptual" side of natural language phenomena that result from interfacing with an expert system.

CHAPTER 4

4.0 Introduction to Expert Systems

An expert system is a class of powerful computer programs whose knowledge domain is focused on a specific task and/or data-base and allows a user to function as if she had keyboard access to a human "expert" in that domain. The expert system has four main components:

1. Input section,
2. Output section,
3. Knowledge base,
4. Access rules.

A user and an expert system exchange information in a stepwise fashion: First, the user types the input to the system. Second, the system translates this input into the command or query language used in its program. Third the system sends an acknowledgment that it understands the input and, if necessary, seeks clarification. Fourth, the system executes the command or outputs the query response to the user.

In this chapter several expert systems are examined for the following features:

1. Knowledge domain,
2. Organization of the knowledge domain,

3. Type of input accepted,
4. Type of output to user.

This information about the expert system's features is summarized and examined relevant to building a NLP. The literature provides more information about how users behave when using computer systems in general and expert systems in particular. These two sources will then be used to specify what features expert systems must provide for in order to interface with a transportable NLP. The expert systems examined are:

1. SCHOLAR
2. INTELLECT
3. KNOBS
4. ISOBAR
5. CRYSLIS
6. PROSPECTOR
7. ESCA SPECTRA INTERPRETER

4.1 The Expert Systems

4.1.1 SCHOLAR

SCHOLAR [Carbonell 1970] is an Intelligent Computer Aided Instruction (ICAI) system whose domain is the geography of South America. It is organized as a semantic

net. A semantic net is a network of information nodes. At each node is a country and the facts associated with that country. Important things to note about this system:

1. It has extensive explanation and text generation capabilities. It does more than just parrot the expert system program back to the user. It is able to ask questions of the user and generate answers which parallel the reasoning of the user. The MYCIN expert system generates answers using "fuzzy logic"; quite unlike the logic of a human user.[Handbook of AI] SCHOLAR's reasoning is similar to that of the human expert and is, therefore, easier to understand.

2. It has extensive inferencing abilities which are used to tailor its instruction to individual students.

3. It has very good focusing abilities. It tracks the topic of discussion with an "agenda". When the discussion becomes ambiguous, it responds by asking questions about what is presently on its agenda.

4. It uses natural language as its input. Using natural language as the input allows the user to respond in the "natural" way of a student: with thought and questions. This also means that the questions will be loosely constructed and possibly ambiguous. When

ambiguities arise, SCHOLAR generates questions about the topic on the agenda, and hopefully puts the session back on track.

A sample dialog (copied from Carbonell) is in Appendix A. It gives a good idea of the kinds of words used and the structure of the utterances. The words as well as the structure are quite simple. Note that SCHOLAR can "understand" what it has written to the user and respond to questions about its replies.

5. The parser uses a case grammar for communicating with the knowledge domain. The parser sends the "domain" the name of the object, and the attribute under discussion. It also sends the "verb" which tells the domain what the user expects the domain to return for and answer. [Carbonell 1970]

4.1.2 INTELLECT

INTELLECT is a portable data-base access expert system. [HARRIS 1979] This system is organized as a data-base query expert system which accepts natural language as input. The dictionary contains the needed closed-class words to converse with any data-base. INTELLECT uses the fields of the data-base to which it is attached for its open-class words. A sample dialogue with

INTELLECT is in Appendix A. Note the short sentences and that they are confined to the data-base under discussion. This points out that even "naive" users are aware that INTELLECT knows only about the data-base it is attached to. They do not make the mistake of thinking that just because they can use natural language with the system, that it can be treated as a "person".

The system has actually been turned over to commercial users and some of the feedback from the users is presented here:

1. There was confusion in the natural language processor about whether an abbreviation was an abbreviation or a word (with a totally different meaning). The example cited was ME as a first person referent or ME as the abbreviation for the state of Maine. As listed in the dialogues in Appendix A, a valid user query might be "PRINT FOR ME THE NAMES OF ALL SECRETARIES". The interpretation, as explained by Harris, could be "INTELLECT, print the names of all the secretaries in the data-base, and do it for me, the user"; or it could be "INTELLECT, print the names of all secretaries in the state of Maine that are listed in the data-base." Under these circumstances, INTELLECT asks questions to find out what you wanted. The point here is that although the interpretation the system may give to a query might seem

obscure to the user, in a given data-base, the system could be justified. Therefore, this is a design point that must be considered when designing a NLP.

2. Another problem INTELLECT surfaced is what Harris called "The Snowmobile Problem". He recommends fitting INTELLECT's natural language processor with ability to compute answers to many different types of "How Many" questions. The examples he cites are: "How many snowmobiles are there?"; "How many secretaries are there?"; and "How many salaries are there?". Each of these how many questions require a different set of calculations to answer. It is plain that these "how many" calculations should be a function of the data-base access expert system and not the NLP itself. (In INTELLECT's case the data-base access expert system is part of the natural language processing expert system.)

3. The third problem encountered was Harris's "Matrix Problem". Sometimes a user requires information from different fields in the data base to answer a question about one item. An example given is, "Give me the total sales for 1972 and 1973." INTELLECT must access the years 1972 and 1973 to find the total sales for each year. This "required combining the syntactic information gained from the parse with domain-specific information about the data base." INTELLECT was not able to do this.

Harris solved this problem "by allowing the specification of how the time referents and the data-base fields to be related." This relation was added to the dictionary. In the case of a NLP the proper place to implement the function of relating data-base fields is in the data-base access expert system.

4. The fourth problem Harris reported on is "The Summary Problem". In the data-bases there is often fields of information which summarize information contained in other fields, such as "total sales for salesman Smith". If the system is not aware that such a total already exists as part of the data-base, it will calculate this total. Again, Harris implemented the solution to this problem in INTELLECT's dictionary.

5. Note that INTELLECT uses the objects in the domain data-base as its open-class words. It passes the objects and their attributes as requested by the user without knowing if the data-base has the information available or even if the object or attribute form part of the data-base.

4.1.3 KNOBS

KNOBS (Knowledge Based System is a "planner" system applied to Tactical Mission Planning. This system accepts

natural language as input and a sample dialog is in Appendix A. The domain is organized as scripts. The system parser passes the conceptual dependency of the objects under discussion to the scripts domain analyzer. The analyzer is a "MYCIN-like" forward-chaining production system. Its purpose is to explore the possibilities of using an AI planning system to help in tactical mission planning. Some of the problems uncovered in KNOBS were due to the limitations of the natural language processor:

1. More dictionary definitions were needed.
2. Better text generation capabilities were needed.

KNOBS used question/answering patterns to generate answers to the user. There were not enough patterns available to fit all the question/answer situations which arose.

3. KNOBS has poor error recovery. When the system could not generate an answer, whether because it did not understand the input or because the data was not in its knowledge base, the system simply failed. The designers think that correcting the above two problems and application of more conceptual constraints and more conceptual completion inferences would make the system more robust. [Engelman, et al 1979] [Milne 1983a]

4.1.4 ISOBAR

ISOBAR (Information Understanding System of Basic Weather Report) produces weather reports based on natural language (Japanese) and pictorial inputs. An interesting feature of this system is its attempt to allow both linguistic and processed pictorial images as inputs. The linguistic portion of the expert system is organized around a restricted semantic grammar. The pictorial data is a digitally coded version of the analog image which is stored separately from the linguistic input. Both input types are then processed and integrated into a third data-base which can be accessed by a user.

The pictorial inputs took too long to process into the data-base and require a complete overhaul to make this a successful method of input. The natural language inputs, however, were a success. ISOBAR accepted a very restricted linguistic input:

1. Declarative sentences - properly formed or else they were rejected,
2. Yes-No questions,
3. WH questions,
4. Commands,
5. 200 word vocabulary.

A sample dialog is in Appendix A. The system can initiate

dialog for clarification of ambiguities or improper grammar using pattern matching. The system's designers have recommended that the natural language capability be more flexible and handle a larger variety of input sentences and allowable words. [Kawaguchi, et al 1979]

4.1.5 CRYSALIS

CRYSALIS identifies the structure of a protein based on information supplied by a "picture" of the electrons of the protein molecule. The structure of this expert system is as a "blackboard" wherein different portions of the system programs can all access and modify the same pocket of data. This data pocket is a hypothesis of what protein is being looked at and this hypothesis is updated depending on what information comes from the different portions of the program. [Engelmore 1979]

This system does not accept natural language as input. This does not mean that it cannot or that natural language input is not needed. Protein crystallography is a complex subject and requires years of training and study to become an expert on. When you spend that kind of time becoming an expert on a subject, you should not be required to become an expert on computer use as well. An expert system on crystallography is well suited for natural language processor use. [Handbook of AI]

4.1.6 PROSPECTOR

PROSPECTOR is a computer-based consultant to help a geologist assess the likelihood of a site containing certain ore deposits. In this respect it is similar to CRYNALIS because they both are for use by experts in their field and not for laymen. PROSPECTOR's contribution to the field of expert systems is its use of both logical and probabalistic reasoning. Geological exploration is as much an art as a science (so say the geologists). Therefore, in this field your logical reasoning about the ore deposits at a given site are only probably correct. PROSPECTOR emulates the reasoning of the geologist. When questions are asked and clarifications are needed, the using geologist can follow PROSPECTOR's reasoning to see if they agree with PROSPECTOR's conclusions.

The system is organized as three models (this number to be expanded in the future) with sets of production rules. The user inputs all the data she has about the site under discussion. PROSPECTOR uses these inputs to decide which model most "probably" fits the site. The productions are then used to decide whether this is correct or not. This approach provides the dialog with focus, ie the model currently under consideration.

The inputs to this system are concise. There is a sample dialog in Appendix A. Most of the user inputs are to state facts to help PROSPECTOR make decisions and to ask "wh" questions. This is a very interactive system because PROSPECTOR is functioning as a consultant to the expert user. Its vocabulary is limited to geological terms and those words needed to make queries about geology.

There are not many linguistic mechanisms built into this system such as ellipsis deciphering. As you will note from the dialog in the appendix, the user and PROSPECTOR managed without them. There was no report from the users that this interaction was too restrictive. However, it should be noted that the users were people who were intimate with the structure of the PROSPECTOR system. It is easy to extend from the experience of the INTELLECT system, that once this system is turned-over to a "real" user, that the human engineering aspects of interaction will become much more important. [Hart 1979] [Gaschnig 1979]

4.1.7 ESCA SPECTRA INTERPRETER

ESCA (Electron Spectroscopy for Chemical Analysis) is an expert system which directly processes spectral data of

a material in order to identify its composition. This program is unlike many other expert systems because it focuses on complex low-level processing instead of high-level inferencing which is the usual fare of AI programs such as MYCIN. The system is organized as a series of productions which analyze the digitized output of a machine called a spectrometer.

The system uses input from a spectrometer which functions by beaming X-rays at a sample material to make some of its electrons jump around. It then records the energies of the jumping electrons. These energies are then used as input to the ESCA Interpreter program. The program, like that of CRYVALIS, is intended to be used by and expert in the field of chemical analysis.

This system does not use natural language as input which is why it is included. The utility of a natural language front-end, as explained in CRYVALIS and demonstrated by PROSPECTOR, is to free the expert of these fields from the task of becoming computer experts as well.
[Yamazaki 1979]

4.2 Requirements Expert Systems Must Meet

The previous section examined the input, output, knowledge domain, and its organization, of seven representative expert systems. This section draws on these characteristics to decide what expert systems must do or provide in order to interface with a NLP. Keep in mind that the NLP should provide the linguistic features, and the expert system the "concept" features.

4.2.1 Vocabulary

There are two aspects of the vocabulary of a NLP/expert system interface: closed-class words and open-class words. The closed-class words come along with the NLP no matter what expert system is used. The open-class words are decided by the expert system. This section examines both the open and closed class words a user needs in order to use the NLP/expert system.

4.2.1.1 Character of the Expert System Vocabulary

In general, the vocabulary of the sample dialogs in Appendix I are typical of the words that people use when using expert systems. Users ask mostly "wh-" questions which means that the most used words support questions of

this type. The verb "to-be" is the most common verb. "You" is the most common subject and the object noun is usually an aspect of the domain of the expert system. The sentences are usually concise. In some cases, ESCA Spectra Interpreter, input is restricted to a specific form and the user is told ahead of time. In LADDER (See Appendix I) the dialogs follow this pattern anyway even though LADDER will accept more than this. However, as also shown by these dialogs, this is not always the case, which is why you need a NLP and cannot make-do with a pattern matcher like ELIZA (described in Chapter 2).

Each dialog contained "domain specific" words which users expect to be there. For example, in the DENDRAL chemical analysis system[Handbook of AI], chemical names such as benzene and methanol, must form part of the vocabulary. These same words are not necessary for systems with a different domain; for example, SOPHIE [Handbook of AI] which teaches electronic troubleshooting.

Another aspect of the vocabulary requirements relates to information about what is included in the domain. Users asked for information about the domain itself, for example: "Tell me about Uruguay". So the system must be able to list information about Uruguay as well as what it knows it knows about Uruguay .

4.2.1.2 Vocabulary requirements

It is a straightforward process to determine what words are needed in the vocabulary of an expert system. The UNIX operating system has a facility which keeps track of all the words that user's misspell most often. This is an indirect methods of keeping track of what words are used most often, not just misspelled. A historical account has to be performed over a reasonable period of time to see what words people use when interfacing with the expert system. [Thomas 1977] These are the words that people need. It is usually the same words used most of the time. Studies were done to find out what would happen if not enough words were provided in the lexicon. [Kelly 1975] Results showed that while people got frustrated, they were still able to perform the task at hand. However, the goal to keep in mind is that you want to make the system as "friendly" as possible and do not need to harass a user who is trying to find oil and make a million dollars with an expert system with the expert system's picky lexicon.

The lexicon must also include phrases. The Air Force lingo is just full of catchy phrases without which the average unit would have to close shop tomorrow. There are two types of phrases that have to be considered; the common type that make up everyday, everyminute language,

and the kind that have a very specific meaning which is divorced from the meanings of the words which make it up, such as "adverse weather delivery system" which is not a contingency plan for the Post Office; it is a military hardware system which delivers armaments. [Becker, 1977]

In natural language dialogue, a mechanism is required which chooses from among the many meanings that a word has. [Cullingform, Pazzani 1981] In a sentence processing system, a complex mechanism, as described by Cullingform for text, is not required. The problem of choosing the correct meaning of a word from its context is handled very well in some existing parsers.

4.2.2 Knowledge Base

Each of the expert systems in the previous section had their own unique knowledge domain and organization. Each system used that organization which best suited the domain and purpose of the expert system. SCHOLAR used a semantic net with a country at each node and the information about that country hierarchically organized beneath it. The KNOBS system was organized as scripts using conceptual dependency. KNOBS is a "Tactical Consultant" and SCHOLAR a ICAI system. Both of these systems used natural language as input. Most expert systems should be able to use natural language for input. If these systems are to

interface to a transportable NLP, the NLP will put out the linguistic information in the same form no matter what expert system it is attached to. The next section describes what the expert system must do in order to interface with the NLP.

4.2.2.1 Requirements On Knowledge Base

First, the expert system must accept the output of the NLP as its own input. This is the logical ordering to be explained in Chapter 4.

Second, the knowledge base of an expert system must be arranged in a manner in which the data can be found and used. If the NLP asks a question of the expert system, the expert system must be able to consult its data-base quickly to see if the answer is in its knowledge base, and then it must be able to pass the answer to the NLP so the NLP can generate a natural language answer to the user.

The types of organization are almost unlimited. Most investigators recommend "natural organization" to coincide with the organization of thought of people. [Ehrenreich 1980] [Kolodner 1978] [Kolodner 1979] [Woods, Brachman 1978] [Poock 1981] [Chen, et al 1982] [Dejong 1982b] Each of these researchers presents their own argument about how to organize memory for efficient retrieval.

The expert system must be able to update its knowledge base. The updates must come through the natural language processor, broken down into the knowledge storage format of the expert system, "integrated" with the rest of the already stored knowledge, and then entered into a place in the knowledge base where it can be found again and used in the future. Davidson and Kaplan [Davidson 1981] point out a method that performs updates to a data-base. They classify updates as simple, "impossible to perform, ambiguous (done in different ways), and as pathological (producing undesirable side effects)". They also point out that users will update the data-base according to their personal view of the data-base. They recommend that the expert system carry a model of the user and adapt the processing of requests according to its "user model". It is also recommended that the whole update process be "transparent" to the user. [Davidson 1979]

When logical inconsistencies of the knowledge base are discovered, there must be a way to correct the inconsistencies. In order for the expert system to retain its "expert" status, there must be a way to reorder conflicts of information. The user must be able to do this through the natural language processor without a whole lot of hassel. The user must know what exactly has

to be changed or else walked through a "change" process by the expert system. All of this exchange is through the natural language processor. This has to be done without the user having to follow the "ripple" of changes that may have to be made in the rest of the data-base in order to accommodate this one change. When changing a knowledge base, there is always the risk that the system was correct in the first place and the user has just introduced an error. [Waltz 1981]

An update of the information base may also be needed due to new information which renders previous information incomplete. The knowledge base must be able to be expanded and the new information incorporated.

Another requirement of the knowledge base is that the system have knowledge about the knowledge-base itself (meta-knowledge). For example, LADDER should know not only that the Kennedy is a ship that is xyz feet long, but also what a ship is and that it floats on water in the ocean. [Linde 1981] [Finin 1977] [Thomas 1977]

4.2.2.2 Clarifying the Dialog - Explanation

Just as it is unreasonable to expect human/human communication to always be clear, it is unreasonable to expect that human/computer communication will be. An

explanation facility must be provided for the inevitable miscommunications. It has been found [Tennant 1980], [Waltz 1981], [Linde 1981] that users often use vague, ambiguous, and incomplete sentences when using expert systems. The system must then be able to query the user to find out what the user wants it to do. When the system answers, often users need to clarify the information it returned in answer to their question, so a question/answer dialog must be accommodated. It is not enough for either the user or system to say, "I don't understand." There must be a give and take of information to find out who doesn't understand what.

Several reasons for a clarification capability are given:

1. Users do not always ask questions. They often make statements which they intend for the system to use in interpreting the next round of questions. In [Waltz 1981:21], the Planes system, a user may state "I'm only interested in A-7's". There must be a facility within the expert system to guide the discourse through questions about A-7's. Some systems such as Planes need this ability more than others which have a more restricted domain, such as Crysalis where the system might get away with assuming that the topic under discussion is the current material sample in the spectrometer.

2. Users ask for information not in the data-base. At the other end of the spectrum, they also ask questions which require that the system make an infinite search. The expert system must guide the dialog so that the user can get some satisfaction. It cannot merely reply, "I can't answer that question" or "I cannot do that". The user needs to know why so he or she can reformulate the query or statement.

3. Users "expect the system to react to unstated goals as evidenced by the user's attempts to make their intentions understood without completely restating their queries." [Cohen 1981:21]

4. The user "expects the system to be aware of her query failures, and more generally, of the failures of her presuppositions. The user expects that the system will ensure that she is not misled by an incorrect assumption." [Cohen 1981:22]

5. Throughout a session, the user must have the feeling that he or she has control over the system. [Linde 1981] [Ehrenreich 1980] "Immediate feedback must be given to user activities". Redundancy and confirmation provide a measure of reassurance to a user by echoing a paraphrase of the user's original request or statement. [Chafe 1980]

Chafe also recommends that the system make frequent reference to the user. A sample dialog was given in his paper which illustrates his point well which I have copied for your convenience:

"My terminal recently told me the following, at 30 steady characters per second:

The "netlpr" command, when executed between computer center machines, now sets the ownership of net queue files correctly so that "netrm" will remove them and they are listed by the "netq" command.

(System dialog is preceeded by **, the rest is user responses)

**Want to know about the "netlpr" command, where you type in "netlpr"?

Sure.

**You can just use it between computer center machines,

OK? Only if you're up here.

Yeah, I know.

**OK. It'll show you who owns net queue files, if you want to know that. You can use "netrm" to get rid of them, and you can get them listed with "netq". That clear?

Yeah."

Methods of meeting this challenge have been proposed:

1. The system keep a model of the user or the user's goal. While this is convenient on the surface, there are many drawbacks to this method. [Norman 1982]

2. A planning mechanism has also been proposed. [McDonald 1983]

3. The system provide the user with a model of the system, instead of the other way around. [Wynn 1980]

4. Another method is an algorithm proposed by de Kleer, [de Kleer, et al 1977] which makes "implicit" information "explicit" so that it can be dealt with.

5. An interactive algorithm which assumes that the user's utterances will be patterned after the system's so that the system can then infer the user's intent easily since it is already familiar with its own assumptions and structure. [Cohen 1981]

6. Hit or miss methods have been implemented wherein the system simply rephrases the user's input in a supposedly unambiguous format that the user can agree to. Similar to that is a method that forms the systems possible interpretations and lets the user choose among them. [Cohen 1981:20]

An example of a simple explanation capability for the MYCIN system is described by Scott. [Scott 1977] MYCIN's "explainer" has access to the production rules of the MYCIN system and can parrot back those rules it is accessing during a session. This is a very basic method and really doesn't explain the production rules themselves, but it is better than nothing. [Cohen 1981]

4.3 Summary

This information contributes toward a decision on how to organize the knowledge base of the expert system. A task oriented dialog takes on a hierarchical form. A data-base oriented dialog has clumps of closely related utterances. [Linde 1981] Defining the orientation of the expert system's dialog is needed to decide how to focus interaction during a user/computer session. The overall character of natural dialogs is that they are vague and ill-specified. [Woods 1981]. If you want to build useable expert systems now, the best approach is that

taken by the present expert systems, that is, the constrained domain. Otherwise the problems become too large and unmanageable.

CHAPTER 5

5.0 The Natural Language Processor - NLP

The NLP does the linguistic processing on both input from the user and output from the expert system. This chapter describes what facilities the NLP must have in order to be a transportable natural language processor to expert systems.

5.1 Input Facilities

To process the user's input, the NLP must have a parser and a lexicon. The next two sections describe their features.

5.1.1 Parser

As explained in Chapter 2, the parser takes the input utterance apart, formats it, and then passes to the expert system. The parser will have to parse both well-formed and ill-formed utterances.

5.1.1.1 Both well and ill-formed utterances:

1. The parser must have a comprehensive grammar which can understand English utterances. It must be able to provide the same interpretation that a person would have assigned to an utterance. User's only use natural language with other persons, therefore, the utterance will always take the same form as if the listener were another person. This could be difficult sometimes because the parser has no "world-knowledge" except that provided by the lexicon. The ROBBIE parser proposes to do exactly this. [Milne 1983] The RUS parser uses both syntactic and semantic knowledge to approximate this.[Harris 1979]

2. It must accept common utterances such as "wh" questions, yes/no questions, commonly recognized declarative sentences, and reasonably well-formed commands. Regarding the commands, it should recognize utterances such as "give me July sales totals" as "print out the July sales totals on the screen for viewing".

3. It must provide the same interpretation on different utterances that mean the same thing but use different words; paraphrases of one another. Related to this is the requirement to provide the same interpretation for an utterance which uses the same words, arranged differently, but still intended to "mean" the same thing

as another utterance.

4. The parse must be done in a reasonable amount of time so that the user still perceives the system as "interactive". If it takes too long, perceptions about the system change from a problem-solving mechanism to a hurdle to be jumped.

5. It must be able to process one AND two sentence utterances. Some users, especially naive users, find it difficult to state their intentions with only one utterance. The parser should be able to deal with the facet of human nature.

6. It should know that it is having a computer/human dialog via a keyboard and be able to handle dialog about the dialog it is having.

7. It must manage the lexicon for both the NLP and the expert system. One focal point is needed to keep the expert system and NLP communicating.

5.1.1.2 Ill-Formed Utterances

The parser must have the following features:

1. Ellipsis resolution,

2. Anaphora resolution,
3. Pronominal references,
4. Spelling correction,
5. Understand poorly constructed sentences,
6. Handle "partial" understanding of inputs,
7. Gracefull error recovery.

These features are common in the dialogs in Appendix A, as well as peppered throughout the many studies referenced in Chapter 2.

Features such as analogy, metaphor and extensive text entry were not used by expert system users. These features would be convenient for a NLP without application limitations. This NLP is restricted just to expert systems and so are not necessary.

5.2.2 Lexicon

The lexicon will be managed by the parser. It must be composed of the open-class as well as closed-class words. The closed-class words will be toted along with the NLP from system to system. The open-class words must be entered into the lexicon everytime the NLP is used with a different expert system. They must be purged if the NLP moves to a different system.

Attributes of the lexicon are:

1. It must contain all the words commonly used in dialogs with expert systems.
2. The words must be kept up to date, and correctly defined.
3. The lexicon must be organized to allow fast, accurate access.
4. Prefixes and suffixes must be included where applicable.

The Air Force (AF) uses many words peculiar to just the AF. These words must be part of the closed-class words which are part of the NLP lexicon and available to the expert systems the NLP interfaces with. Many common AF words have been compiled and are included in Appendix II.

5.3 Output - Text Generation

It is important to first state that the text being generated in this system is one sentence at a time - not one sentence after another to form a paragraph with a cohesive idea behind it. (That kind of work is much more complicated, but if you are interested, DanLos gives a quick review of how to do it.) [DanLos 1983]

Text generation differs from an "explanation capability" in that the latter outputs a concept and the former puts those concepts or ideas into sentences for direct viewing by a user. Text generation is more like anti-parsing. Text generation and explanation must work hand in hand. The ideas behind the "anti-parse" (if you will) come from the expert system. The words come from the lexicon for use by the "text generator". The PHRED and PHRAN systems of the University of California use this idea. [Jacobs 1983]

It is important that the NLP "understand" what it has just said to the user. The user is likely to ask questions about this output and will expect that the NLP and expert system know what was output.

5.4 Interfacing NLP With The Expert System

The NLP must be modified everytime it is interfaced with a different expert system. The NLP must "understand" the user's request and free it from such phenomena as ellipsis, poor construction, bad spelling and all the other confusing aspects of natural language. A variety of methods of performing this task by the processor have been studied and implemented in different host systems. The natural language processor must be able to provide to the expert system, in a fixed and expected format, what data

is required. This implies that a "development" must take place before hooking the natural language processor to an individual expert system to define the communication format. This is akin to how the INTELLECT system was fielded. INTELLECT was "modified" to interface with each user's data-base by the INTELLECT system designers, tests were run, and then the system was turned-over to the user.

Note that there is no facility for adapting the NLP to the expert system by the domain expert. It is intended that the NLP expert do the actual work. This is not to be confused with extending the lexicon which is intended to be a task the user or domain expert can do.

A standard input/output format is proposed for the interface between the natural language processor and the expert system to which it is attached. One format proposed by [Milne 1983] and others is a logical form; this being most adaptable on a wide scale. An example given by Milne is:

John lives in a green tent.

which resolves as: Lives <John, tent-1> and Color <tent-1, green>.

The expert system then uses this as input to be

resolved using the rules of predicate calculus. He also points out that this form is readily adaptable to the "command, argument" form making this a very flexible format. It is recommended that all future expert systems be able to accept such an input format.

5.5 Summary

This chapter looked at the requirements on the NLP. Chapter 6 will summarize the requirements on both the NLP and the expert systems to which it may interface, as well as future possibilities for the ES/NLP interface.

CHAPTER 6

6.0 Summary

The next two sections summarize the features of both the NLP and the expert systems which may work with with it. Then additional features for the future will be discussed.

6.1 The Natural Language Processor

The following features must be provided by a NLP for expert systems which may be used by the Air Force.

1. Process one and two sentence English utterances from a user.
2. Process utterances that most people would understand if they read them in context. This includes well and ill-formed utterances.
3. Resolve pronominal references, ellipsis, and anaphora.
4. Error handling facility to resolve lexical ambiguity of input.

5. Generate output text for its error handling facility, the expert system domain answers, and the clarification facility of the expert system.
6. Manage a lexicon with words and phrases that most users need in the given expert system domain.
7. Allow user to upgrade the lexicon.
8. Spelling correction.
9. Reasonable processing times for each input.
10. Output a fixed, logical format to be used as input to all expert systems.

6.2 The Expert System

The following features should be included in any expert system which is to be interfaced with an NLP for Air Force expert systems:

1. Knowledge base organization which can accept a fixed, logical format from a modular NLP and be resolved using the rules of predicate calculus.
2. Knowledge base which "knows" facts about its knowledge

domain as well as what it doesn't know about its knowledge domain.

3. Knowledge about what is not included as part of its knowledge domain.

4. Access rules which support the NLP to communicate this knowledge.

5. Detection of logical inconsistencies and ambiguities.

6. Question/Answer facility for clarifications.

7. Reasonable processing times for each input and output.

6.3 What Happens When The NLP/Expert System Hookup Leaks

It would be nice to say that all the specifications of the NLP/ES interface will be laid on the NLP or on the ES. This is not always in the best interests of efficiency or compatible with reality. The TED system [Hendrix, Lewis 1981] allows for the development of "interface information" when attaching a transportable NLP to an expert system. TED's mechanism for adding interface

information allows the user to perform this task. A mechanism should be provided for adding and modifying system/interface information in case the need arises.

6.2 Add-On Features OR Better Technology Through Living

Features such as understanding meaning, and multi-sentence text generation are presently under study. If and when researchers find it practical to improve the technology and capabilities of ES interfaces, this feature provides a convenient place to put it.

Features which allow processing of visual information are also envisioned. Visual representations of the expert system's knowledge domain greatly expand the utility of a system which interfaces with users who do visual manipulations.

Appendix A

This appendix contains actual dialogues user's have had with each expert system. The user's utterances are preceded by a double asterisk (**). An analysis of the dialogs' content is in Chapter 4.

A.1 SCHOLAR

Start

What is the language of Argentina:

**probably Spanish

Very good.

Approx what is the area of Argentina?

**Tell me something about Peru

Peru is a country.

It is located in South America.

The capital of Peru is Lima.

Now answer the question you didn't answer before.

**The area of Argentina is 1,100,000 square miles

Approximately true, but a more exact value is 1,200,000.

The eastern bordering country of Peru is --

**Catholicism

Wrong. Catholicism is not a country but a religion.

The correct answer is: Brazil.

Where is Peru?

** Is Asuncion hot?

Since Asuncion is part of Paraguay, yes.

**What is the most important difference between Peru and Paraguay?

The primary difference is: The population of Peru is larger by a factor of 7.8.

End SCHOLAR

[Handbook of Artificial Intelligence]

A.2 WHY

Start

Do you think the Amazon jungle has heavy rainfall or light rainfall?

****heavy rainfall**

Why does the Amazon jungle have heavy rainfall?

****because it's near the Atlantic**

yes, the Amazon jungle is near the equatorial Atlantic.

How does that affect rainfall in the Amazon jungle?

****The water in the current is warm**

What happens after the moisture-filled air reaches the land?

****I don't know**

Do you think the Andes mountain range has any effect on the rainfall in the Amazon jungle?

****yes**

How?

****The prevailing wind hits the Andes and the air cools**

How does the fact that prevailing wind hits the Andes, lead to air cooling?

Do you think that any place with mountains has heavy rainfall?

****yes**

Southern California has mountains. Why doesn't Southern California have heavy rainfall?

End WHY

[Handbook of Artificial Intelligence]

A.3 SOPHIE

Start

Welcome to Sophie, a reactive learning environment. The piece of equipment you are studying is a regulated power supply. Would you like to start with a faulted instrument that needs fixing?

****yes**

The instrument has a fault and its present control settings are:

. .blah

.

.blah

.

.blah

.

****what is the output?**

The output voltage is 1.7 volts.

****what is it in a working instrument?**

In a working circuit the output voltage is 19.9 volts.

****what is the voltage between R8 and Q1?**

The voltage at N1 is 47.6 volts.

****Replace Q5**

I am going to ask you some questions about how Q6 is faulted.

Are any junctions shorted?

****i don't know**

Are any terminals open?

****yes**

Which terminal(s) are open?

****the base**

Q6 is not faulted in that manner. Q6 not replaced.

****is it possible that Q6 has an open base?**

Is the complete fault mode EB/EC/BC = Open/Ok/Open what you intended?

****sure**

That fault is consistent with all of the measurements you have made.

End SCHOLAR

A.4 KNOBS

Start

**WHAT IS THE HARDNESS OF TARGET 2?

THE HARDNESS OF TARGET 2 IS UNKNOWN

**HOW ABOUT TARGET 1?

TARGET 1 IS VERY HARD

**WHY?

BY TR3: TARGET 1 IS VERY HARD

SINCE:

1; DATA: TARGET 1 IS A BRIDGE

** IS TARGET 3 AN AIRBASE?

NO, TARGET 3 IS AN SA-4

**CHOOSE MUNITIONS FOR IT

THE BEST MUNITIONS FOR TARGET 3 IS SHRIKE

** WHY?

BY MR4: THE BEST MUNITIONS FOR TARGET 3 IS SHRIKE

SINCE:

1: BY TR1: TARGET 3 IS AN EMITTER

SINCE:

1: Y TR4: TARGET 3 IS A SAM

SINCE:

1: DATA: TARGET 3 IS AN SA-4

(See Jane's 1978, pg 77 and 150 for an explanation of what a SHRIKE AND SA-4 is.)

End KNOBS

A.5 ISOBAR

Start

This dialog is translated from the Japanese. Its value lies in the words which make up a dialog between an user and a weather reporting expert system.

**Was the weather on the side of the Japan Sea almost fair at noon on Dec. 30, 1978?

Yes, the weather on the side of the Japan Sea was almost fair.

**In NAHA area at noon on Dec. 12, 1977, was the weather south?

The weather has not the attribute of direction.

**Where did a warm front extend from the center of the low pressure at noon on Sep. 12, 1976?

A warm front extended to N26deg E140deg from the center of the low pressure.

**When did the pressure configuration of typical winter type occur?

At 6:00p.m. on Nov. 29, 1976, the pressure configuration of typical winter type occurred around Japan.

End ISOBAR

A.6 INTELLECT

The following are actual queries put to this system. The system answers are not shown:

Start

**SHOW SUBTOTALS OF DIRECT COMMISSION BY MONTH FOR 1ST QTR 78 IN REGION M WHERE NET ARISING AMOUNT IS AT LEAST \$150,000.

**FOR DOMESTIC MACHINERY, PRINT THE JULY 1979 SALES DOLLARS AND SALES QUANTITIES, SORTED BY SALES CLASS.

**FOR COMMERCIAL DIVISION AND THOSE PRODUCTS IN EXCESS OF \$500,000 YEAR-TO-DATE, SHOW ME ALL TIME PERIODS WHOSE PRE-TAX EARNINGS ARE MORE THAN \$200,000 AND PRINT THEIR NAME.

**FOR SUMMARIES WITH LOSS RATIO GREATER THAN 200, REPORT REGION, BRANCH AND LOSS RESERVE BY MONTH.

**PRINT ALL PRODUCTS WHOSE NOVEMBER 1977 SALES QUANTITIES ARE BETWEEN 100,000 AND 300,000 AND WHOSE PRE-TAX EARNINGS EXCEED \$1,000,000.

****PRINT FOR ME THE NAMES OF ALL SECRETARIES**

End INTELLECT

A.7 PROSPECTOR

This system uses simple English for dialog. The session has the following phases and are invoked by typing the word for that phase:

? Prints a paraphrase of a given question
WHY Prints the reason that PROSPECTOR asked the question
SUMMARIZE Prints a summary of conclusions
VOL Allows user to volunteer information to PROSPECTOR
CHANGE Allows user to change an answer to a previous question
COMMENT Allows comments to be added to transcript file
WAIT Interrupts session in order to modify the program
QUIT Stops program execution
HELP Enters a "help" routine

Start

****THERE ARE CARBONATES**

CARBONATES (4)

PROSPECTOR acknowledges the first volunteered evidence: the certainty is taken to be 4.

****THERE IS DOLOMITE**

DOLOMITE (4)

****THERE IS GALENA**

GALANA (4)

****THERE MIGHT BE SPHALERITE**

SPHALERITE (2)

****DONE**

(The user is finished volunteering evidence, and PROSPECTOR takes the initiative.)

Models are (PCDA MSD MVTD)

(These are abbreviations of the three current models.)

You told me about: CARBONATES (4)

To what degree do you believe that:
Prospective ore body lies in carbonate sediments?

**3

You told me about: GALENA (4)
You suspected: SPHALERITE (2)
Your statements imply: SULFIDES (4)

To what degree do you believe that: Prospective ore body
contains galena and
spalerite?

**2

You told me about: DOLOMITE (4) - CARBONTES (4)
To what degree do you believe that: Prospective ore body
involves dolomite?

**?

(The user responds to the question with "?" and a
clarification of the question is furnished.)

Assuming that there is an ore body in carbonate sediments,
to what degree do you believe that some significant
dolomitization of the host carbonates has occurred? Note:
the exact type of dolomitization is not important.

**3

To what degree do you believe that: Fluid inclusion data
suggest that mineralization took place above 160 degrees
C?

**0

To what degree do you believe that: Prospective ore body
contains non-telethermal alteration products?

**4

To what degree do you believe that: Other evidence
indicates that mineralization took place above 160 degrees
C?

**0

To what degree do you believe that: Evidence of metamorphism?

**WHY

The evidence of high temperature mineralization is discouraging for the prospective ore body being a MVT. However, if the high temperatures were due to subsequent metamorphism, then this discouraging evidence should be discounted or ignored. To what degree do you believe that: Evidence of metamorphism?

**4

Your statements imply: Sulfides (4)

To what degree do you believe that: Prospect contains sulfide minerals other than those of Fe, Pb and Zn?

**3

To what degree do you believe that: Any prospect assay indicates more than one ounce per ton of silver?

**0

To what degree do you believe that: Prospect contains fragmental host rock?

**5

To what degree do you believe that: Ore mineralization is stratabound?

**0

To what degree do you believe that: An unconformity or disconformity, or several of them, overlie the prospective ore body?

**0

To what degree do you believe that: Favorable stratigraphic sequence for a Mississippi Valley type deposit?

**?

The stratigraphic sequence almost always found with a MVT is as follows: blah..blah..blah... Ignoring faulting, erosion, or subsequent deposition of material, to what degree do you believe that this

describes the original stratigraphic sequence?

**2

My certainty in suggestive evidence for MVTD is now:
4.997

Do you want to pursue this further?

**yes

You told me about: CARBONATES (4)

To what degree do you believe that: Carbonate sediments containing the prospective ore body are shallow-water shelf-type sediments deposited under stable tectonic conditions?

**3

To what degree do you believe that: Presence of hydrocarbons?

**5

This dialog goes on until PROSPECTOR summarizes its interpretation and concludes with:

I cannot do any more with this hypothesis, Do you want to pursue another one?

**QUIT

End PROSPECTOR

A.8 LADDER

LADDER prints out information about making connections with the network, invoking IDA (now called SODA) which I have left out because it is not of interest in the context of natural language processing.

Start

**What is the current position of the Kennedy?

POSITION 6000N03000W DATE 7601171200

This means that the Kennedy was last reported to be at 60 degrees North, 30 degrees West, at noon on January 17, 1976.

**of kitty hawk

Trying Ellipsis: WHAT IS THE CURRENT POSITION OF KITTY HAWK

LADDER informs the user of all of its actions. Sometimes this is important, for example, when LADDER has to reinterpret a question, the user should know what the question is that he or she is getting an answer to.

POSITION 3700N01700E DATE 7601171200

**To what country does each merchant ship in the north atlantic belong?

I have 79 answers; do you want to see them all?

**yes

NAT	SHIP	POSITION	DATE
US	GREENVILLE VICTORY	1533N03010W	7601171200
US	TOULE J	1533N03010W	7601171200
	.		
	.		
	.		
	.		
WG	TRANSQUEBEC	4633N02129W	7601171200
NO	TALABOT	0600N01130W	7601171200

** the fastest sub

Trying Ellipsis: TO WHAT COUNTRY DOES THE FASTEST SUB BELONG

NAT US SHIP LOS ANGELES MXPD 30.0

**what ships are en route to america

Trying Ellipsis: Ellipsis has failed.
The parser does not expect the word "AMERICA" to follow
"WHAT SHIP ARE EN ROUTE TO"
Options for next word or symbol are:
COUNTRY PORT
end-of-list

LIFER cannot parse the question;
"America" is not in the vocabulary.

**what is a "country"

In addition to questions about the data base, INLAND can answer questions

about the grammar itself. The need for this capability in all expert systems is discussed in Chapter 4.

COUNTRY may be any member of the set {ANGOLA ANGOLAN ARABIA ARABIAN ARGENTINA ARGENTINAN BRITAIN BRITISH CANADA CANADIAN DUTCH EGYPT EGYPTIAN ENGLAND ENGLISH FOREIGN FRANCE FRENCH GERMAN GERMANY HMS H.M.S. ITALIAN ITALY LIBERIA LIBERIAN NETHERLANDS NORWAY NORWEGIAN PORTUGAL PORTUGUESE RUSSIA RUSSIAN SOVIET SPAIN SPANISH U.D. U.S. U.S.A. U.S.S U.S.S.R. UK US USA USS USSR VENEZUELA VENEZUELAN}
finished

****define america like usa**

(The user may add new synonyms to the vocabulary. There is danger as well as convenience in this ability. The convenience is that you can add needed definitions on the spot when the need arises. The danger is that the user might not know the proper use of a given word and may be adding an improperly defined word to the vocabulary.)

FINISHED

REDO Redo is a feature of Interlisp.
SHIP = KENNEDY JF, LOS ANGELES, BATON ROUGE, PHILADELPHIA, POGY, ASPRO, SUNFISH, KAWISHIWI

(Notice the use of punctuation. While very simple here, people have grown used to its presence.)

****how many of them are navy ships**

(Them or she is currently always interpreted as a reference to a set of ships in the previous query. A portable natural language interface cannot be allowed to do that.)

8 of them:

SHIP = KENNEDY JF, LOS ANGELES, BATON ROUGE, PHILADELPHIA, POGY, ASPRO, SUNFISH, KAWISHIWI

****give status kitty hawk**

Trying Ellipsis: Ellipsis has failed.
The parser does not expect the word "STATUS" to follow "GIVE". Options for next word or symbol are:
RELATIVE.CLAUSE, SHIP, VALUE.SPEC, THE
end-of-list

**define (give status kitty hawk) like (list the employment schedule, state of readiness, commanding officer and position of kitty hawk).

EMPLMNT SURVOPS EMPBEG 760103 EMPEND 760205 READY 2 RANK
CAPT NAME SPRUANCE R POSITION 3700N01700E DATE 7601171200

**give status us cruisers in the mediteranean

spelling => MEDITERRANEAN

EMPLMNT:	CARESC	CARESC	CARESC
CARESC			
EMPBEG:	760101	751231	751231
751231			
EMPEND:	760601	760615	760615
760615			
READY:	1	1	1
1			
RANK:	CAPT	CAPT	CAPT
CAPT			
NAME:	MORRIS R	HARMS J	EVANS O
FRENZINGER T			
POSITION:	4000N00600E	3700N01700E	
3700N01700E	3700N017003		
DATE:	7601171200	7601171200	7601171200
7601171200			
SHIP:	CALIFORNIA	DANIELS J	WAINWRIGHT
JOUETT			

**done

Thank you.

End LADDER

Appendix B

Vocabulary

ablating materials
ablation
access authorization
access time
accident scene
accident site
accompanied baggage
acquisition cycle
active element
actual range
additional flying training periods
administrative action
advanced development
adverse weather aerial delivery system
advisory airspace
advisory signal
advocacy
advocacy package
aerial combat tactics
aerial mining
aerial survey team
aeroballistic missile
aerospace ground equipment
aerospace medicine
aerospace security force
aerospace systems security program
agravic
Air Almanac
airborne direct air support center
airborne launch control center
airborne launch control system
airborne radio direction finding
airborne stationkeeping
airborne warning and control system
air cargo
aircraft arresting hook cable tie-down
aircraft arresting retractable hook cable
aircraft commander
aircraft fault diagnosis
aircraft hook cable support
aircraft integrated data systems
aircraft marshalling
aircraft operating weight
aircraft parking
aircraft possessed
aircraft salvage

aircraft structural integrity program
aircraft turn around
aircrew authorized
aircrew available
aircrew formed
air distance
air drag
airdrop
air eligible cargo
Air Force Budgeting and Accounting System for Operations
Air Force Central Membership Association Fund
Air Force central welfare fund
Air Force-controlled unit
Air Force industrial fund
Air Force in-house laboratory
Air Force integrated command and control system
Air Force recoverable assembly management system
Air Force Reserve (Associate) Unit
air freight terminal
air intercept missile
airintercept rocket
airlift command post
airlift control center
airlift control element
airlift force
air line of communications
airman advanced training
airman basic technical training
airman lateral training
airman promotion and test scoring data system
airman supplemental training
air mileage indicator
air movement designator
air munitions
Air National Guard
Air National Guard of the United States
Air operations center
air plot wind velocity
Air Reserve Forces
Air Reserve Forces Policy Committee
Air Reserve technician
air run landing
air run takeoff
air speed indicator
air taxiing
air technician
air traffic advisory service
air traffic control line
air traffic control system
air traffic regulation center
air transportation housekeeping equipment
air transportable loading dock
air weapons control system

Allied Communication Publication
allotment
altimeter
altimeter setting
altitude azimuth
altitude circle
altitude difference
altitude/height hole
altitude intercept
analog to digital converter
analog computer
angle of jump
angular velocity sight
aphelion
apparent altitude
apparent precession (apparent wander)
apparent solar day
apparent solar year
apparent wander
application study
approved program
arc of fire
arc tunnel
area defense
area monitoring
area of probability
area of uncertainty
arithmetic unit
armed conflict
arming signal
Army reconnaissance liaison officer
arresting net stancion
arresting sheave
arresting sheave span
arresting system cycle time
arresting system payout
arresting system purchase element
arresting system reset time
arresting system reset unit
arresting system runout
aspect angle (angle off)
assault airdrop
assault airland
assembler (data automation)
assessment center
associate contractor
assumed position
assured destruction
astrionics
astrodome
astronomical triangle
astro procomputation
atmospheric braking

atmospheric refraction
atomic time
attack sortie
attitude jets
audiovisual
audiovisual communication
audiovisual product
aural null
authority
authorized data chain
authorized procurement data list
automated data processing system
automated data system
automatic approach and landing
automatic data processing
automatic data processing auxiliary equipment
automatic data processing equipment
automatic data processing maintenance services
automatic data processing product
automatic data processing program
automatic data processing system
automatic direction finder
automatic flight control system
automatic test equipment
autumnal equinox
availability rate
average wind velocity
azimuth
azimuthal projection

backup interceptor control system
ballistic coefficient
ballistic reentry
ballistics
band of error
bare base
barrier spacing
base level military personnel system
base level self-sufficiency spares
base line extension (navigation)
base line (navigation)
base self-sufficiency
basic military training
beam
beam-width error
bellamy drift
below threshold changes
bench check
bench test
bioenvironmental engineering
bio-pak
Bird Dog
black body

blind toss
blip scan
bomb sighting systems
boresight line
boundary security
briefing directive
bubble horizon
budget authorization
budgeting cycle
buffer
built-in test equipment
bulk cargo

calculated altitude
calibration card
camouflage detection photography
captain (aircraft commander)
captive test
card programmed
career assistance and counseling unit
career development
career information and counseling unit
career management
casual uplift
caution signal
caveat
celestial azimuth
celestial equator
celestial horizon
celestial poles
celestial precomputation
central security control
checkout equipment
checkout time
chief controller
chief interceptor controller
circle of equal altitude
circle of position
circular error
Civil Air Patrol
clinic
close-in security
close security area
cocked hat
cockpit familiarization trainer
cockpit procedures trainer
cold mission
collateral damage
combat center
combat hamlet
command and control communications
command destruct
command ejection system

command select ejection system
common-user communications
communication language (data automation)
communications (signal center)
compass acceleration error
compass calibration
compass swing
component force
component life - indefinite
component life - our of phase
compression of the earth
computed altitude
computing gunsight
conceptual phase
cone of dispersion
cone of fire
confidence test
configuration
conformal projection
consolidated base personnel office
constellation
container delivery system
contamination control
contamination line
continental United States commands
contingency air terminal
continuation military training
continuously computed impact point
continuously computed release point
continuously computed target point
continuously set vector
continuous strip imagery
controlled area
controlled general way
conversion (data automation)
co-operative scorer
coordination
copy negative
copy plot
correctional custody
correction or change, projected reassignment action
corrosion
Corsair II A-7A
cost attribution
cost avoidance savings
cost effectiveness
cost plus incentive fee contract
countervalue
course training standard
crew duty time
crew ratio
crew training
crippled mode

critical engine failure speed
critical field length
critical point
cross trail
cross trail angle
curve of equal bearing

dagger
damaged shipment
damage expectancy
damage limitation
dangerous material
data automation
data automation requirement
data base
data conversion equipment
data descriptor
data display
data handling
data link equipment
data management officer
data name
data processing
data processing center
data processing installation
data project directive
data reduction
data system
data system automation program
data system designator
data system interface
data systems integration
data system specification
data transcription equipment
data transmission equipment
data use identifier
dead reckoning
debug (data automation)
decayed object
decca
deck alert
declaration of excess
decoy
defended hamlet
deflection component of trail
deflection error
deflection error probable
degaussing
delayed opening chaff
demilitarize
demonstration and shakedown operations
density altitude
departure

AD-A138 455

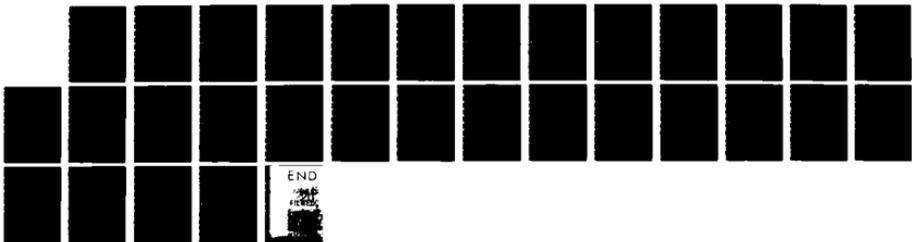
NATURAL LANGUAGE PROCESSOR AS A UNIVERSAL FRONT END TO
EXPERT SYSTEMS(U) AIR FORCE INST OF TECH
WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING
D L BROOKE DEC 83 AFIT/GE/EE/83D-13

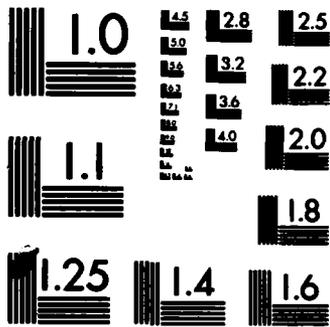
2/2

UNCLASSIFIED

F/G 9/2

NL





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

deployment phase
depression
deputate
designated system
design studies
design to cost
desired heading
destruct line
detail group
detail part
detection
development
development concept paper
development planning studies
development project
development testing and evaluation
D factor
diagnostic routine
digital computer
digital to analog converter
dip
direct access
direct indicating compass
directional guroscope
direction center
direction finding course
direction finding net
direction finding penetration and approach procedures
Directorate of Automatic Data Processing Equipment
Selection
director sight
disaster cordon
discs
dispatch interval
dispatch rate
dispersed operating base
distortion (lens)
dive angle
diversionary missile
dive toss
document
documentation (data automation)
dog let
donuts
doppler radar
dosage
dose rate
double rail logic
downwash
Draft Presidential Memoranda
drift
drift meter
drogue chute

dropping angle
drop zone control officer
D sounding
dual base concept
dual thrust

earth satellite
ecliptic
effective atmosphere
effective miss distance
electron
electronic
electronically controlled photography
electronic defense evaluation
electronic element
electro-optical countermeasures
electro-optical guidance
emergency combat capability
emergency exposure limit
emergency security operations
emergency war order
emergency war order alert
emissivity
employment
end system resolution
energy effectiveness
enroute base
enroute support team
entry controls
envionics
ephemeris time
equation of time
equipment augmentation (data automation)
equisignal zone
escape velocity
essential material
estimate
Executive Manpower Resources Board
ex-meridian altitude
exotic fuel
experimental model
extended active duty
externally-programmed automatic test equipment
extraction zone
extraction zone control officer

false colour film
false horizon
familiarization job training
familiarization training
fault
fiber optics
field extension

field of view
field of vision
field training
field training detachment
filled allocation
film base (photograph)
filter
final heading
Financial Assistance Programs
first generation negative
first generation positive
first pilot
First Point of Aries
First Point of Libra
fiscal guidance
Five Year Defense Program
fixed reticle
flag stop
flight
flight facilities
flight information publications
flight path
flight plan fuel load
flight profile
flight simulator
floating reticle
fluxgate
fluxgate compass
fluxvalve
flying crane
focal point
folded optics
follow-on developmental tests
follow-on operational test and evaluation
Force and Financial Program
formal logic
formal training
forward operating base
forward recovery mission profile
forward supply activity
forward supply point
forward supply support
free gyroscope
frequency band
fuel-air explosive munitions
full-scale development phase
functional test

general military training
general mobilization
generation
generation (photography)
geocentric latitude

geodetic latitude
G-force
geographic latitude
global commands
go-no-go
go-no-go check list
go-no-o test equipment
gravity anomaly
gravity drop
gravity drop angle
Greenwich hour angle
Greenwich sidereal time
grid navigation
gross altitude scale
Gg air vehicle
Gg alert
Gg cushion
Gg half-coupling
Gg performance aircraft
Gg petroleum products
Gg resolution
Gg support equipment
Ggwash
gunbore line
gun jump
gunsight line
gunsight radius
gyro-graph
gyro horizon
gyro log
gyro-magnetic compass
gyroplane

harmonize
hard savings
Headquarters Air Force
Headquarters Air Force Reserve
head up display
helicopter
hiran
home
home recovery mission profile
hot mission
human engineering
hush house
hyperbolic navigation system

ideal bomb
identification
image degradation
imagery collateral
imagery correlation
imagery data recording

imagery exploitation
imagery pack
impact point
inactive status
inadvertent release
incomplete round
independent check
independent ejection system
index error
index register
indicated air temperature
indicated altitude above mean sea level
indicated pressure altitude
individual development plan
individual training
induced precession
industrial preparedness measures
industrial production base
information program
infrared countermeasures
infrared detector
infrared film
infrared linescan
infrared radiation
infrared reconnaissance
infrared resolution
infrared sensors
inherent distortion
initial heading
initial operational test and evaluation
input
input/output
instantaneous field of view
instream factor
instruction
instructional systems development
instrumentation
instrument flight trainer
integrated communications navigation and identification
integrated weapon support management
intelligence information handling systems
intercardinal points
intercept
interceptor missile
interference
intermediate staging base
internal controls
interpretability
intersite communications
intertheater airlift
intransite communications
intertheater airlift
intransit air freight

intrasite communications
intratheater airlift
intrusion detection equipment
inverted stereo
involuntary active duty

jam-resistant data link
job training standards
Joint Air Training Command Facility
joint task force
jump (data automation)
jump speed

kill effects
kinetic lead
knot

large outsize cargo
laser
laser guided bomb
laser illumination
laser linescan
laser ranging
laser reconnaissance system
laser target designation
lashing point
lashings
lateral axis
launch
launch vehicle
lead angle
leading line
Lead-the-Force Program
levy
L-Hour
life cycle cost
life cycle costing
life cycle cost procurement
life support research
light armed reconnaissance aircraft
lighter-than-air aircraft
light filter
light intratheater transport
line astern
line of aim
line of constant bearing
line of departure
line of position
line of sight
load category
load designator
local apparent noon
local hour angle

local sidereal time
logic
logistic and administrative support
logistic planning
longitudinal axis
lost shipment
low altitude parachute extraction system
lower limit
low light level television

machine oriented language
magnetic bearing
magnetic compass
magnetic field
magnetometer
main stage
maintainability
maintenance check
major installation
major subsystem
management analysis
management and support
management engineering
maneuvering ballistic reentry vehicle
man-hour availability
man-hour utilization
manpower availability
manpower utilization
manual test equipment
master caution signal
master station
master warning signal
materiel handling equipment
materiel repair schedule
maximum aircraft hook load
maximum short takeoff and landing weight
maximum/vertical takeoff and landing weight
mean solar day
mean area of effectiveness
mean time between failure
mean time to repair
mechanized materiel handling system
medical center
medium high combat environment
memory
meridian altitude
meridian passage
military accepts responsibility for separation of aircraft
military assistance service fund
military service obligation
military standard item characteristics coding structure
military standard transaction reporting and accounting
procedures

military supply and transportation evaluation procedures
millibar
minimal flight path
minimum fuel
minimum lethal dose (dosage)
minimum line of detection
minimum line of interception
miss distance scorer
missile procedures trainer
missile system
mission degradation factor
mission trainer
mobile training assistance
mobile training set
mobilization production planning
mobilization production requirement
modified close control
modified precision approach radar
modular repair/overhaul
most probable position
movement control team
multi-drift wind velocity
multiple ejection rack
multiple independent reentry vehicle
multiple options
multi-spectral imagery
munitions
munitions delivery error
munitions delivery error probable

national defense area
National Military Command Center
national range
negative altitude
negative G
negotiation thresholds
night observation device
non-co-operative scorer
nondestructive inspection
non-operating active aircraft
non- US citizen
northerly turning error
notional tasking
notional unit
not operationally ready, maintenance
not operationally ready, supply
not reparable this station
nuclear, biological and chemical defense
nuclear defense
biological defence
chemical defence

octant altitude

office of collateral responsibility
office of primary responsibility
off-line
off-line storage
omnirange
on-condition maintenance
on-line
onload base
operand
operating active aircraft
operating command
operating ground equipment
operating location
operating weight
operational aircraft damage/loss
operational concept
operational element manning requirement
operational readiness training
operational status
operational suitability
operational support
operational test and evaluation
optical countermeasures
optical instrumentation
organizational mobility equipment set
outside cargo
oversea commands
oversea duty
over shipment

parallax in altitude
parts kit
part task trainer
passive element
penetration area
perihelion
periodic inspections
periodic maintenance/inspection
permanent change of station
permanent sensor installation
permissive reassignment
persistence
personal error
personnel monitoring
personnel subsystem
photographic flight line
photo processing cell
pipper
pitch
pitot-static tube
pitot tube
point defense
pop-up maneuver

positive control of nuclear weapons
positive G
powered lift flight regime
power plant unit
precedence rating
precession
precession of the equinox
precision approach radar
prediction angle
preliminary warning notice
preposition war reserve materiel
Presidential Memoranda
pressure altimeter
pressure altitude variation
pressure line of position
pressure pattern navigation
prestocked war reserve materiel
prevailing visibility
preventive perimeter
primary alerting system
primary inspection
primary radar
prime airlift
prime BEEF
prime meridian
priority designator
program management responsibility transfer
professional military education
proficiency training
program advocacy
program change decision
program change request
program element
program element monitor
program evaluation and review technique
programmable stimuli
program management directive
program management plan
programmed base
program memorandum
programmer comparator
projectile velocity
project officer
proportional navigation
protected system
provisioned spares
proximity scorer
pseudo stereo
pucks
pure pursuit course
pursuit course interception
pylon

Q correction
quadrantal points
quality control
quick engine change unit

radar boresight line
radar prediction
radar ranging
radar scan
radial
radial error
radial velocity
radiant energy
radiological survey flight altitude
range component of cross trail
range error
range error probable
rapid area distribution support
Rapid Engineer Deployment, Heavy Operational Repair
Squadrons, Engineering
rated supplement
Ready Reserve agreement
real precession
real property installed equipment
real time
real wander
reclamation
reconnaissance pallet
reconnaissance pod
reconnaissance reference point
reconnaissance slipper
reconnaissance system turnaround
recovery base
rectified altitude
recycle base
reentry system
reference meridian
refire time
reflex sight
relative wind velocity
remote indicating compass
repair cycle
replacement training unit
report control symbol
report, general, mission review
second phase
report, immediate
report, in-flight
reporting point
report, photographic interpretation
reports management system
required operational capability
research coupling

reseau
Reserve of the Air Force
resolving power
retaliation
reticle
retractable runway edge sheave
retrograde cargo
retorocket
reversal film
reversal process
reverse stereo
right ascension
ring sight
rocket ammunition
rocket artillery
rocket assisted take-off unit
rocket engine
rocket motor
rocketry
rocket sled
rocket thrust
rockeye
roll control
rolling vertical takeoff
rollover
rolometer
rotor
rotorcraft
route transport operations
round

safe haven
S-bend distortion
scan line
scan rate
S-curve distortion
sea launched ballistic missile detection and warning
system
second pilot
secure data link
secure visual communications
security police forces
security reporting/alerting system
security subsystem
seeker
self-contained night attack
self-test
semiautomatic ground environment
semiautomatic test equipment
sensible atmosphere
sensible horizon
sepration
sequenced ejection system

service engineering
service test model
sextant altitude
shake-table test
shiran
short range attack missile
short shipment
Shrike
sidereal day
sidereal hour angle
sidereal time
sidereal year
sighting angle
sight tracking line
signal to noise ration
silo
simulated-operations testing
sky compass
sky wave
sky wave correction
LORAN
snap report
solid propellant binder
sonic boom
sounding rocket
sound suppressor
special air mission
special assignment airlift mission
special technical instruction
specialty training standard
specific tasking
spherical triangle
split-altitude profile
split base concept
spot hover
stadia metric ranging
standard air munitions package
standard arm
standard time
standoff weapon
state of the art
static firing
static marking
static pressure
station set
status analysis
step-climb profile
stick
strategic airlift
strategic attack
strategic hamlet
strategic nuclear weapon
strike control and reconnaissance

strip
strobe marker
subassemble
substrate
support base
support equipment
support system
survivability
swinging compass
symbolic logic
synchronous orbit
system justification statement
system life cycle
system phasing
system program director
system program office
system/project key subordinates
system management
system safety
system safety engineering
system security engineering
systems engineering
systems management
Systems Source Selection Board
system turnover

tabulated altitude
tachometric sights
synchronous sights
tactical air control party
tactical air control system
tactical aircraft shelters
tactical airlift
tactical electronic warfare
tactical fighter replacement squadron
tactical fighter training squadron
takeoff weight
tear down
temporary change of station
thermal crossover
thermal resolution
thermal shadow
threshold limit value
tie-down fittings
time base
time change items
time compliance technical order
time lag
time over target
tongue width
total active inventory
total aircraft authorization
Total Force Policy

total inactive inventory
total inventory
total obligational authority
trace
track and ground speed wind velocity
tracking index
track made good
traffic control and landing systems
trail angle
trail distance
trail formation
train bombing
trainers
training priority category system
trajectory
trajectory scorer
trajectory shift
trample
transceiver
transit bearing
transition envelope
transition flight
transmitter-receiver
triangle of velocities
triple ejection rack
tropical year
trouble shooting
true azimuth
turbofan
types of training courses

unaccompanied baggage
Uniform Materiel Movement and Issue Priority System
unilateral function
unique data system
United States Air Force Reserve
United States Air Force Special Operations
United States Air Force Special Operations Forces
unit equipment
unit training assembly
unit training standard
universal time 0
universal time 1
universal time 2
universal time co-ordinated
unnatural hostile environment
unsatisfactory report
unscheduled maintenance
upper limit
upward mobility
urgency of need designator
utilidor
utilization rate

validation phase
variable geometry aircraft
vector
vector gunsight
vector sights
velocity jump
vernal equinox
vertical and short take-off and landing capability
vertical circle
vertical launch
violated shipment
visible horizon
visual report
VPR

Walleye
war consumables
warning receiver
warning signal
war readiness spares kit
war reserve materiel
war termination
weapon delivery
weaponeering
weapon system performance recording
weapon systems physical security
weapon system trainer
weather cocking
whole range distance
whole range point
windage jump
wind factor
wind triangle
word length
workweek

Z-correction
zenith
zenith distance
zero gravity
zone time

MAJCOM
avionics
comptroller
auditor general
Settlement and Adjudication
Acquisition Logistics
B-1 Program
Blue Flag & Exercise Evaluation
Biodynamics & Bioengineering
Applications & Liaison Office

Consent Judgement
Civilian Institutions
Command & Control
Combat Crew Training (Survival)
Chief of Staff
Combat Crew Training (Aircraft)
Vice Commander
Commandant of Cadets
AFIT
Dean of Faculty
Civil Engineering
Operations
Personnel
Distribution
Headquarters Commandant
TAC
SAC
MAC
USAFE
PACAF
OSD
OSI
Chaplain
Green Flag
First Sergeant
Aide to the Commander
Commander
Protocol
Unit OJT
A-10
A-7
T-37
T-38
F-16
F-15
B-1
B-52
EF-111
AFLC
Remotely Piloted Vehicles (RPV)
Pentagon
Chief of Staff
Vice Chief of Staff
Inspector General
NATO
Fighter Pilot
Line Officer
Rated Officer
Non-Rated Officer
Under Secretary
World Wide Military Command & Control System
USSR
TEMPEST

Warsaw Pact
antiaircraft artillery (AAA)
antiair warfare (AAW)
antiballistic missile (ABM)
US Navy AAW weapon system (AEGIS)
artillery-fired atomic projectiles (AFAP)
Air Force Satellite Communications System (AFSATCOM)
air-launched cruise missile (ALCM)
airborne launch control system (ALCS)
advanced medium-range air-to-air missile (ASRAAM)
long-range phased-array radar (AN/FPS-85)
SLBM detection and warning radar (AN/FSS-7)
Alternate National Military Command Center (ANMCC)
Association of South East Asian Nations (ASEAN)
ASW rocket (ASROC)
antisubmarine warfare (ASW)
antitank
advanced technology bomber (ATB)
All Volunteer Force (AVF)
Airborne Warning and Control System (AWACS)
ballistic missile defense (BMD)
Soviet airborne amphibious combat vehicle (BMD)
Ballistic Missile Early Warning System (BMEWS)
command and control
command, control, and communications
command, control, and communications countermeasures
close air support
Combined Forces Command (CFC)
CONUS ground station
CHAMPUS
comprehensive test ban
Defense Communications System (DCS)
distant early warning
Division Air Defense
Defense Logistics Agency (DLA)
Demilitarized Zone (DMZ)
electronic counter-countermeasures (ECCM)
electronic countermeasures (ECM)
electronic intelligence (ELINT)
Emergency Rocket Communication System (ERCS)
electronic warfare (EW)
foreign military sales (FMS)
ground-launched cruise missile (GLCM)
gross national product (GNP)
general purpose forces (GPF)
intercontinental ballistic missile (ICBM)
imagery intelligence (IMINT)
intermediate-range nuclear forces
initial operational capability (IOC)
intermediate-range ballistic missile (IRBM)
individual ready reserve (IRR)
Committee for State Security of the USSR (KGB)
Atlantic Area (LANT)

line of communication (LOC)
long-range theater nuclear forces (LRTNF)
Marine Amphibious Brigade (MAB)
Military Airlift Command (MAC)
Marine Amphibious Force (MAF)
Mobile Command Center (MCC)
multiple independently targetable reentry vehicle (MIRV)
Multiple Launch Rocket System (MLRS)
medium-range ballistic missile (MRBM)
multiple rocket launchers (MRL)
noncommissioned officer (NC)0
North American Aerospace Defense Command (NORAD)
over-the-horizon-backscatter (OTH-B)
Perimeter Acquisition Radar Attack Characterization System
(PARCS)
Precision Acquisition of Vehicle Entry and
Phased Array Warning System (PAVE PAWS)
permanent change of station (PCS)
Palestine Liberation Organization (PLO)
program objective memorandums (POM)
People's Republic of China (PRC)
reentry vehicle (RV)
radar warning receivers (RWR)
Strategic Air Command (SAC)
Strategic Arms Limitation Talks (SALT II)
surfact-to-air missile (SAM)
single integrated operational plan (SIOP)
submarine-launched ballistic missile (SLBM)

sea-launched cruise missile (SLCM)
short-range nuclear forces
short-range attack missile
Strategic Arms Reduction Talks (START)
tactical fighter wing (TFW)
TOMAHAWK Land Attack Missile-Nuclear
Theater of Military Operations
United Nations (UN)
United States Army (USA)
United States Marine Corps (USMC)
United States Army Reserve (USAR)
United States Marine Corps Reserve (USMCR)
United States Merchant Marine (USMM)
United States Navy (USN)
United States Naval Reserve (USNR)
Union of Soviet Socialist Republics

Bibliography

Becker, Joseph D. "The Phrasal Lexicon", Proceedings of the Fifth International Joint Conference on Artificial Intelligence , Tokyo, Japan 1977

Birnbaum, Lawrence and Mallory Selfridge. "Problems in Conceptual Analysis of Natural Language", ARPA Research Report No 168, October 1979.

Bobrow, Robert. "The RUS System: Research in Natural Language Understanding", BBN Report NO. 3878, Cambridge, MA: Bolt, Beranek and Newman, Inc.

Brown, David R., Darrell V. Fowler, David J. Frie, William M. Hadly, Joan C. Introne, William T. Park, Ann E. Robinson, William Schubert. "R & D Plan For Army Applications Of AI/Robotics", SRI Report Number ETL-0296, DTIC No. Ad-A118277, May 1982.

Carbonell, Jamie R. "AI in CAI: An artificial intelligence approach to computer-aided instruction", IEEE Transactions on Man-machine Systems, MMS-11(4), 190-202, 1970.

Chafe, Wallace L. "Should Computers Write Spoken Language?", Report on the Workshop on Interactive Man-Machine Discourse held at the University of Pennsylvania June 17-20, 1980, DTIC No. AD A096512.

Chien, R. T., Bill Frederick, Michael Houghton, Adam Pajerski, "Diagnosis and Design of Mechanisms using Deep-Level Understanding Models", July 1982, DTIC No. AD A120184.

Cohen, Philip, Raymond C. Perrault, James F. Allen, "Beyond Question-Answering", May 1981, DTIC No. AD A100432>.

Cullingford, R. E. and J.J. Pazzani, "Word Meaning Selection in Multiprocess Language Understanding Programs", 1981, University of Connecticut, paper available DTIC.

DanLos, Laurence. "Some Issues In Generation From A Semantic Representation", Proceedings of the Eighth International Joint Conference on Artificial Intelligence 1983.

Davidson, Jim and S. Jerrold Kaplan, "Natural Language Access to Databases: Interpreting Update Requests", DTIC No. AD A109183.

DeJong, Gerald F. and David L. Waltz, "Understanding Novel Language", July 1982, DTIC No AD A120184.

DeJong, Gerald F. and Paul O'Rourke, "Knowledge Based Automatic Schema Acquisiton", July 1982, DTIC No. AD A120184.

de Kleer, Johan, Jon Doyle, Guy L. Steele, Jr., Gerald Jay Sussman, "Explicit Control of Reasoning", June 1977, DTIC No. AD A078244.

Ehrenreich, S. L. "Design Recommendations for Query Languages", U.S. Army Research Institute for the Behavioral and Social Sciences Technical Report 484, DTIC No. AD A115874.

Engelman, E. Miriam Bischoff. "KNOBS: An Experimental Knowledge Based Tactical Air Mission Planning System and a Rule Based Aircraft Identification Simulation Facility", Proceedings of the Sixth International Joint Conference on Artificial Intelligence Tokyo, Japan 1979.

Engelmore, Robert and Allan Terry, "Structure and Function of the Crysalis System", Proceedings of the Sixth International Joint Conference on Artificial Intelligence , Tokyo, Japan 1979.

Finin, Tim, Bradley Goodman, Harry Tennant, "JETS: Achieving Completeness Through Coverage and Closure", Proceedings of the Fifth International Joint Conference on Artificial Intelligence 1977.

Gaschnig, John. "Preliminary Performance Analysis of the Prospector Consultant System for Mineral Exploration", Proceedings of the Fifth International Joint Conference on Artificial Intelligence 1977.

Grosz, Barbara. "TEAM: A Transportable Natural-Language Interface System", SRI Technical Note 263R, November 1982>

Harris, Larry R. Proceedings of the Sixth International Joint Conference on Artificial Intelligence, Tokyo, Japan 1979.

Hart, P. E. and R.O. Duda "PROSPECTOR -- A Computer-Based Consultation System for Inneral Exploration" SRI Technical Note No. 155.

Hayes, Philip J. and Raj Reddy, "Flexible Parsing", October 12, 1981, DTIC No AD A109076.

Hayes, Philip J. "Anaphora for Limited Domain Systems", Available DTIC and ARPA Order 3597, 1981.

Hendrix, Gary G. and William H. Lewis, "Transportable Natural-Language Interfaces to Databases", SRI Technical Note 228, 30 April 1981.

Jacobs, Paul S. "Generation in a Natural Language Interface", Proceedings of the Eighth International Joint Conference on Artificial Intelligence 1983.

Kawaguchi, Eiji, Masao Yokota, Tsutomu Endo, Tuneo Tamati, "An Understanding System of Natural Language and Pictorial Pattern in the World of Weather Report", Proceedings of the Sixth International Joint Conference on Artificial Intelligence, Tokyo, Japan 1979.

Kelly, M. J. "Studies in Interactive Communication: Limited Vocabulary Natural Language Dialogue", Johns Hopkins University Technical Report 3, August 1975, DTIC No. AD A019198.

Kolodner, Janet L. "Memory Organization for Natural Language Data-Base Inquiry", Technical report #142 for Advanced Research Projects Agency of the DOD, 1978.

Kolodner, Janet Lynne. "Retrieval and Organizational Strategies in Conceptual Memory: A Computer Model", Technical report #187 for Advanced Research Projects Agency, DTIC No. AD A093084, 1979.

Lehnert, Wendy. "Affect Units and Narrative Summarization", DTIC No. AD A086735, May 1980.

Levin, Beth C. "Instrumental With and the Control Relation in English", DTIC No. AD A078420, November 1979.

McDonald, Proceedings of the Eighth International Joint Conference on Artificial Intelligence 1983.

Miller, George A. and Philip N. Johnson-Laird, "Language & Perception", The Belknap Press of Harvard University Press, Cambridge MA., 1976.

Miller, Lance A. and Curtis A. Becker "Programming In Natural English", IBM Technical Report No RC5137, November 15, 1974>

Milne, Robert. "Natural Language Understanding For Military Systems", Proceedings for Conference of American Institute of Aeronautics and Astronautics Computers in Aerospace No. IV , Hartford, Connecticut, October 1983.

Milne, Robert. "Resolving Lexical Ambiguity in a Deterministic Parser", PhD. thesis, Department of Artificial Intelligence, University of Edinburgh, 1983.

Norman, Donald A. "Some Observations on Mental Models", DTIC No. AD A116031, 1982.

Rowley, M. B. "Foundations for the Development of a Simple Natural Language Interface for Task Knowledge Elicitation and Representation", AMTE(E) TM 81102, AMTE (Teddington), Teddington Middlesex TW11 OLN, Copyright Controller HMSO London 1982.

Scott, A. Carlisle, Randall Davis, William Clancey, Edward Shortliffe, "Explanation Capabilities of Production-Based Consultation Systems", Stanford University Report Number STAN-CS-77-593, Available DTIC, February 1977.

Schank, Roger C., Michael Lebowitz, Lawrence A. Birnbaum,
"Integrated Partial Parsing", DTIC No. AD A062638,
November 1978.

Schank, Roger C. and Lawrence Birnbaum, "Memory, Meaning,
and Syntax", DTIC No. AD A095451, November 1980.

Tennant, Harry Ralph. "Evaluation of Natural Language
Processors", PhD Thesis University of Illinois at
Urbana-Champaign, Available DTIC, 1981.

Thomas, John C. "Quantifiers and Question-Asking", IBM
report No. RC5866, February 1976.

Thomas, John C. "A Method for Studying Natural Language
Dialogue", IBM Technical Report No. 5882, DTIC No. AD
A056121, July 1977.

United States Military Posture , Prepared by The
Organization of the Joint Chiefs of Staff, Fiscal
Year 1983.

Waltz, David L. "On the Interdependence of Language and
Perception", Proceedings of the Sixth International
Joint Conference on Artificial Intelligence , Tokyo,
Japan 1979.

Webber, Bonnie Lynn. "A Formal Approach to Discourse Anaphora", BBN Report No 3761, DTIC No. AD A055671, May 1978.

William A. and R.J. Brachman "Research In Natural Language Understanding BBN Technical Report #3742 for Office of Naval Research ARPA order #3414, 1978.

Winston, Patrick Henry, Artificial Intelligence, Addison-Wesley Publishing Company, Reading, Mass., 1979.

Woods, W.A. "Procedural Semantics As A Theory of Meaning", BBN Report No. 4627, DTIC No. AD A097630, 1981.

Wynn, Eleanor. "What Discourse Features Aren't Needed in On-Line Dialogue", Report on the Workshop on Interactive Man-Machine Discourse, University of Pennsylvania, DTIC No. AD A096512, June 1980.

Yamazaki, Masato and Hideo Ihara, "Knowledge-Driven Interpretation of ESCA Spectra", Proceedings of the Sixth International Joint Conference on Artificial Intelligence, Tokyo, Japan 1979.

VITA

Captain Deborah L. Brooke was born on 27 September 1952 in Baltimore, Maryland. She graduated from high school in Chittenango, New York in 1970 and attended the State University of New York at Stony Brook from which she received the degree of Bachelor of Science in Electrical Engineering in May 1978. Upon graduation, she received a commission in the USAF through OTS. She was stationed at Hanscom AFB, Massachusetts until entering the School of Engineering, Air Force Institute of Technology, in June 1982.

Permanent address: 6 Grove Place
Danbury, Connecticut

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFIT/GE/EE/83D-13		5. MONITORING ORGANIZATION REPORT NUMBER(S)	
6a. NAME OF PERFORMING ORGANIZATION School of Engineering	6b. OFFICE SYMBOL (If applicable) AFIT/ENG	7a. NAME OF MONITORING ORGANIZATION	
6c. ADDRESS (City, State and ZIP Code) Air Force Institute of Technology Wright-Patterson AFB, Ohio 45433		7b. ADDRESS (City, State and ZIP Code)	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	
8c. ADDRESS (City, State and ZIP Code)		10. SOURCE OF FUNDING NOS.	
		PROGRAM ELEMENT NO.	PROJECT NO.
		TASK NO.	WORK UNIT NO.
11. TITLE (Include Security Classification) See box 19			
12. PERSONAL AUTHOR(S) Deborah L. Brooke, B.S., Capt, USAF			
13a. TYPE OF REPORT MS Thesis	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day) 1983 December	15. PAGE COUNT 126
16. SUPPLEMENTARY NOTATION Approved for public release: LNW AFR 190-17. 2 Feb 84 Lynn E. WOLFE Dean for Research and Professional Development Wright-Patterson AFB OH 45433			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	Artificial Intelligence, Natural Language	
09	02	Expert Systems, Computational Linguistics	
19. ABSTRACT (Continue on reverse if necessary and identify by block number)			
Title: Natural Language Processor as a Universal Front End to Expert Systems			
Thesis Chairman: Robert W. Milne, Capt., USA			
Abstract: This thesis specifies a list of features which must be included in a transportable natural language processor (NLP) used as a front-end to expert systems used by Air Force personnel. It also places requirements on the expert systems themselves in order to interface with such a NLP.			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL Robert W. Milne, Capt., USA	22b. TELEPHONE NUMBER (Include Area Code) 513-255-3576	22c. OFFICE SYMBOL AFIT/ENG	

END

FILMED

384

DTIC