AD-A083 923 NAVAL OCEAN SYSTEMS CENTER SAN DIEGO CA F/G 5/7
NATURAL LANGUAGE PROCESSING APPLIED TO NAVY TACTICAL MESSAGES.(U)
FEB 80 D M KEIRSEY
UNCLASSIFIED NOSC/TD-324 NL

1 OF 1
AD
A083923

NOSC

END
DATE
FILMED
-80
DTIC

1 LEVEL II

NOSC

NOSC TD 324

NOSC TD 324

Technical Document 324

NATURAL LANGUAGE PROCESSING APPLIED TO NAVY TACTICAL MESSAGES

Tactical Command and Control Division
(NOSC Code 824)

David M Keirsey
(Systems Development Corporation)

1 February 1980

Prepared for
Naval Electronic Systems Command
(Code 330)

Approved for public release; distribution unlimited

NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CALIFORNIA 92152

NAVAL OCEAN SYSTEMS CENTER, SAN DIEGO, CA 92152

AN ACTIVITY OF THE NAVAL MATERIAL COMMAND

SL GUILLE, CAPT, USN
Commander

HL BLOOD
Technical Director

ADMINISTRATIVE INFORMATION

This work was performed by the Tactical Command and Control Division (Code 824) as part of the C3 Systems Theory problem under NOSC Project No. CC25. The project is sponsored by NAVELEX, Code 330, the Command and Control Division of the Research and Technology Directorate, NAVELEX, Code 03. Work was performed by Mr. David M. Keirsey of Systems Development Corporation under Contract N00123-78-C-0190, Task 012.

Released by
RC Kolb, Head
Tactical Command and
Control Division

Under authority of
JH Maynard, Head
Command Control–Electronic
Warfare Systems and
Technology Department

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NOSC Technical Document 324 (TD 324)	2. GOVT ACCESSION NO.: AD-A083 923
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): NATURAL LANGUAGE PROCESSING APPLIED TO NAVY TACTICAL MESSAGES	5. TYPE OF REPORT & PERIOD COVERED
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Tactical Command and Control Division (NOSC Code 824); David M. Keirsey (Systems Development Corporation)	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Ocean Systems Center, San Diego, California 92152	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61153N XR01408, XR0140801
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Electronic Systems Command (Code 330)	12. REPORT DATE: 1 February 1980
	13. NUMBER OF PAGES: 60
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

(14) NOSC/TD-324

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Information systems	Message processing
Intelligence–Artificial	Natural language
Machine translation	Tactical communications

Accession For: NTIS GRA&I ✓ / DDC TAB / Unannounced / Justification / By / Distribution/ / Availability Codes / Dist: A / Avail and/or special

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Investigates Natural Language Processing (NLP) as a possible solution to the problem of automating the fusion of multisource data, and finds that if there is a solution to the problem of timely use of text-like data in the Naval tactical message, then that solution is NLP. However, major modifications to the "human" side of the process are needed; for example, uniform usage of abbreviations and acronyms. Tactical messages are analyzed. The problems are addressed in terms of NLP techniques. The relevant NLP techniques are described. Sample "knowledge representations" are proposed. It is concluded that NLP is an infant technology, and technology transfer from NLP and related efforts elsewhere will be difficult. Great care is needed in developing a usable, working system. Care in specifying exactly the capabilities and deficiencies of the system is important.

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Table of Contents

OBJECTIVE

The military decision maker is unable to effectively use, in a timely manner, the increasing amount of diverse data available. The purpose of this task is to investigate Natural Language Processing (NLP) to determine if it has useful application to the problem of automating the fusion of multisource data.

RESULTS

1. If there exists a solution to the problem of timely use of text-like data in the Naval tactical message, then that solution is certainly NLP. However, some major modifications to the "human" side of the process are needed. For example, a uniform usage of abbreviations and acronyms would greatly simplify the problem.

2. Tactical messages are analyzed. The problems are addressed in terms of NLP techniques. The relevant NLP techniques are described. Sample "knowledge representations" are proposed.

3. It is concluded that NLP is an infant technology, and technology transfer from NLP and related efforts elsewhere will be difficult. Great care is needed in developing a usable, working system. Care in specifying exactly the capabilities and deficiencies of the system is important.

RECOMMENDATIONS

1. An experimental NLP system should be designed and built.

2. An expert in Naval operations, someone who understands the content of the messages, should be included in the project.

3. A study on the design or modification of future information systems should be conducted by a team of people. The composition of that team should include experts in naval tactical operations, information systems, and artificial intelligence.

1. INTRODUCTION

1.1 Scope

This is a report of the work done in trying to apply Natural Language Processing (NLP) techniques to the "free format" text in the Navy's communication network messages. Navy tactical messages of a partly formatted type have been studied extensively to determine the major problems and the possible success that NLP might have in solving them. This work is part of the larger problem of "automated data fusion" [5].

1.2 The Problem

The messages contain tactical data that is distributed to various parts of the network. The information transmitted includes ship positions, weather reports, contacts, and results of missions. The messages of concern here are primarily in "fixed format." That is, the information is encoded into fixed fields that are predefined. For example:

AREA/4600N6/16500E3/100NM

This specifies an area. However, sometimes accompanying the fixed-format lines are comments. These comments can be in any form --- free format. The free-format lines are typically amplification on the fixed-format information. The comments vary from a few words to 20-30 sentences. The problem is to try to incorporate the fixed-format and free-format information into an automated data acquisition and analysis system. The fixed-format text can be understood by a machine, whereas the free-format text can only be understood by people. A possible solution is to apply NLP techniques to the comments.

The comments, Naval tactical messages (NTM), have several special characteristics compared to written prose. Some characteristics make the problem more difficult and others make it easier.

The messages deal with a narrow range of activity (narrow compared to arbitrary prose), which makes the problem of understanding easier. Successful AI systems (MYCIN, DENDRAL, INTERNIST, HEARSAY) rely heavily on limited problem domains. The envisioned system will first deal with a subset of messages. The growth of the system will be dependent on the ability to understand the various activities of the Navy. This document will concentrate on messages that report contact with unknown platforms. Only simple events such as searching, tracking, and detecting will be dealt with.

1.3 An Example

The following is representative of typical free-format text.

> NARR/ 2 FLARES SIGHTED 230704Z6 SOUTH. APPROX 5 MI SPA ESTABLISHED. CENTER SPA 226 K HAWK 9 MI. INVESTIGATED AREA. CONTACT GAINED 230735Z3. CLASSIFIED POSSUB CONFIDENCE 3 TRACKING SOUTHEAST SPEED 17. CONDUCTED 2 ASROC ATTACKS. LOST CONTACT 230743Z9.

The fixed-format text associated with the above told of a contact with a submarine, and a position fix was established.

There are several properties that are exhibited by this and other messages.

1. Sentences are not grammatical; they are incomplete.
2. The person or thing that did the action is not referred to explicitly in the sentence.
3. Abbreviations are used extensively.
4. The tense is primarily past tense.
5. An enormous amount of acronyms and jargon is used.
6. Sentences are frequently run together (missing period).
7. Misspellings occur occasionally (more than normal prose).
8. Standard tactical situations are described.
9. The fixed-format text associated with the free-format text gives a good indication of the possible content.

2. OVERVIEW OF NATURAL LANGUAGE UNDERSTANDING

2.1 Why NLP?

It is possible that some useful information is contained in the narrative part of a message. The problem then becomes: to get the information in a form usable by various reasoning programs, such as STAMMER [1] or HASP [reference available to qualified requesters]. Take an example sentence:

> Surface contact turned away upon approach and began stern chase.

By the rules of grammar, this sentence is incorrect; the subject of the sentence is not the subject of the phrase "began stern chase." Despite this, people understand the sentence. A more explicit description might be:

> The surface contact turned away upon our approach.
> We then began a stern chase.

Any number of useful facts could be gleaned from this sentence. They might be: two platforms changed course, the surface contact is hostile (inferred), or the platforms are now on the same course (inferred by "stern chase"). All these facts could be useful. NLP techniques give the machine the ability to "understand" as people do and deduce useful information. Most sentences in the NTM are syntactically ill-formed, or, more precisely, incomplete. This presents a difficulty. Syntax of a language plays an important part in giving structure to a sentence. Without much structure, it is difficult for humans to understand what is being said:

Center spa 226 k hawk 9 mi.

as opposed to:

<u>The</u> center <u>of</u> <u>the</u> spa <u>is</u> 9 mi <u>from</u> <u>the</u> k hawk <u>at</u> <u>a</u> bearing <u>of</u> 226 degrees.

With syntax we can make some reasonable guess to the meaning despite not knowing the words:

The glub of the slith is 9 qr from the kuth.

Despite the fact that syntax is important, the problem is still approachable. The major part of understanding is semantics. For example, take the following sentence from a pen pal from a foreign country.

I suffer very. It reason. I English poor.

In this instance one can ask what she could have said. Clearly her English is poor. She could feel bad about it. Although the syntax is violated, one should understand the meaning. Because we have knowledge of what could have been said, the meaning, "I feel bad, because my English is poor," is fairly evident.

Having some idea what is going to be said is the major key to understanding natural language. Syntax is only one source of knowledge to help understand language.

2.2 The Understanding Approach

The understanding approach to NLP is based on the observation that a great deal of knowledge is needed to parse sentences. Previous approaches in NLP had difficulty with ambiguity. However, people rarely see ambiguity where the machine does. Context and knowledge help to choose the right interpretation of an ambiguous sentence. For instance, the sentence "Contact gained" could either mean: 1) a communication was established; 2) a platform was sighted; 3) a platform, which was chasing the message sender's platform, got closer. When reading the sentence in context (see example 1.3), only one interpretation comes to mind; only one meaning is seen because the context gives some expectation of what could happen.

In another example (sentence six in example 1.3), the correct parsing is difficult without knowledge. The ambiguity is whether "confidence 3" modifies "tracking" or "classified." (The author, not knowing the meaning at first, made the incorrect choice initially.)

2.2.1 The importance of verbs

The verb is the center of understanding in the sentence. Example:

1. Kitty-hawk was flibbed.
versus
2. Gorch was damaged.

Given these two examples, one can see that although both sentences have nonsense words, sentence two is "more understandable." That is, one has more of an idea what happened in two. To make this clearer, the following are some variations of the two examples with real words substituted. One understands sentence 2 better than 1 because an inference can be drawn without knowing the object. E.g. the object's use (or function) has been inhibited. No useful inference can be made if the verb is not known. The following sentences show a greater variation in possible meanings in sentence 1 than sentence 2.

1a. Kitty-hawk was destroyed.
1b. Kitty-hawk was detected.
1c. Kitty-hawk was recalled.

2a. Bomb-1a was damaged.
2b. Plane-1b was damaged
2c. Radar-1c was damaged.

Since the verb is the center of understanding, case grammars [6] are very useful in parsing. The main NLP method proposed in this report is the primitive approach, which derives its foundation from case grammar [10]. The primitive approach relies heavily on finding the semantic relationship between the noun phrases and the verb. Syntax is used to help establish those relationships.

2.2.2 Parsing versus Understanding

In the early part of NLP most of the research was on parsing. Parsing is the process of breaking up the sentence into its syntactic components. Understanding, on the other hand, is not primarily concerned with how the sentence is constructed. The primary concern is to extract the intending meaning in any manner possible. Thus, the understanding process includes parsing as a "knowledge source," a "knowledge source" being information that can help the process. There are multiple sources in the understanding process. Morphological analysis and expectations from the fixed-format text serve as other sources.

2.2.3 ATN Grammars

A fairly successful technique for NLP has been Augmented Transition Networks (ATN) [16]. The basic idea of the ATN is to represent the grammar of the language in a graph-like structure. Parsing with an ATN consists of traversing the network while building up syntactic and semantic structures along the way. The arcs in the graph represent legal terminals and non-terminals of a phrase structure grammar. The computer implementation of this technique is extremely simple. The sophistication is in the designing the grammar and deciding on the semantic structures to build. The basic problem with the technique is that ungrammatical sentences will not parse. ATN's suffer from an all or nothing acceptance of sentences; there is no partial understanding. It is possible to take into account the syntactic errors by defining them as part of the grammar [13]. But in the case of the NTM, the syntactic variation is too difficult for an ATN parser to handle. (It might be better to say, there is not enough structure for the ATN grammar to parse.)

A semantic grammar [3] is another variant on an ATN. A semantic grammar uses semantic predicates on the arcs as opposed to syntactic predicates. This technique also is not considered useful in understanding the NTM. The semantic grammar, like the ATN, must anticipate all possibilities, which is nearly impossible in the NTM. A very flexible scheme for parsing is needed to cope with the variety and unexpected form of the NTM.

2.2.4 Knowledge Structures

With the realization that knowledge is the most integral part of understanding, the development of knowledge structures has been a major thrust in NLP. The knowledge structures "frame," "script," and "plan" are very similar in concept. The notions of semantic network and primitives have also been widely used. This report will view the use of several knowledge structures in conjunction as necessary in understanding the NTM.

2.2.4.1 Primitives

The primitive serves as a building block for representing the rest of the knowledge structures. They represent basic concepts which can serve to cover all possible situations. There was a debate among NLP researchers on what and how many primitives are needed; however, it now is generally agreed that the issue was not important [14]. Any reasonable set can serve as a basis. In this report a variant of Schank's set of primitives will be used. An important function of the primitives is to cover the important features of the Naval tactical situation. The following are deemed to be the important facts.

Contacts
Attacks
Identifications
Movements
Location of objects
Status of objects
Time of events

The primitives are organized into knowledge domains, a knowledge domain being a set of related concepts. Seven domains are seen as necessary to represent the events in the NTM: space, motion, resource, vision, force, goal, and knowledge. Time is a concept which pervades all other domains; it does not have its own domain.

The form of a primitive is a case format. The primitive name is followed by named slots. Example:

2 flares sighted 230704Z6 South.

```
(ATTEND (ACTOR ms) (OBJECT flare) (DIRECTION (BEAR 180))
        (SENSOR EYES))
```

The primitive act ATTEND represents all types of "seeing," e.g., visual sighting, radar sighting, sonar detection, MAD detection, etc. The slots represent the objects and relations to the primitive. The ACTOR slot represents the person(platform) that does the ATTENDing. In most instances the actor will be the message-sender. The OBJECT slot represents the object being seen; SENSOR, the type of detection; DIRECTION, the direction the object is located in; TIME, the time of the act of ATTENDing.

The primitive gives the power of inferring other information.

For example, if (ATTEND (OBJECT flare) (DIRECTION *D1) (TIME *T1)), then a platform is probably in the direction of *D1 at time *T1.

2.2.4.2 Situations

There exist events, at a conceptual level, higher than primitive acts and states. There are combinations of primitive acts that represent some stereotypical situation. It is envisioned that this level of representation will be the main work-horse for codifying the knowledge needed in understanding the NTM. For example, a typical naval situation is one platform attacking another. One might represent ATTACK situation as simply

```
(CAUSE
    (DO (ACTOR ms))
    (CAUSE
      (PROPEL (ACTOR specific) (OBJECT (WEAPON weapon1))
         (DIRECTION (TO contact1) (FROM ms)))
      (PTRANS (ACTOR *) (OBJECT (WEAPON weapon1))
         (TO contact) (FROM ms))))
```

The paraphrase of this representation is: The message-sender did something to cause some force applied to a weapon which causes it to move toward the contact.

However, this representation is missing an important part of the notion attack. Take the situation: John accidentally shot Bill. One would not classify this situation as an attack. But in the simple version of attack, the two situations would be represented the same. The missing part of attack is the goal of the attacker to do harm to the victim. Thus, a situation is a set of related primitive acts and states. Secondly, the situation might be instantiated by different acts. Specifically, the dropping of depth charges or the shooting of guns are different acts, but they are both methods of attack. Lastly, situations predict other acts or situations; e.g., ATTACK predicts either an EVADE or DAMAGED.

2.2.4.3 Scripts

A lot of work has been done by Roger Schank [11] and his colleagues on the notion of scripts. Very much like situations, the script represents stereotypical events. The number of situations in the NTM is much larger than the number of scripts. The difference is that the script represents a set of events that have some type of sequential ordering. An example is a photo mission. The typical photo mission involves a sequence of five actions:

1. A reconnaissance platform is tasked to take pictures.
2. It goes to the intended location.
3. It takes the pictures.
4. It returns to base.
5. It gives its information to the command.

2.2.4.4 Higher Knowledge Structures

Schank has proposed knowledge structures such as "plans" and "themes." They organize knowledge more abstract than primitives, situations, and scripts. They are necessary for general understanding, but they have little use for understanding the NTM. These structures might be applicable to naval messages not tactical in nature, but those messages cannot be tackled with the current NLP techniques.

2.2.4.5 Semantic Networks

The semantic network is one of the oldest ideas in NLP. It also is one of the most overused or meaningless terms in NLP. A directed graph with labeled arcs and nodes, where the arcs are interpreted as relations on the nodes, is the most simplistic notion of a semantic network. The more structured semantic nets restrict the number of possible labels and justify the choice [2].

One important feature of semantic nets, which is difficult in other representations, is the ability to represent hierarchical information. Two relations are needed: Instance and Include. E.g., SHIP is an INSTANCE of PLATFORM, PLATFORM INCLUDES SHIP. A taxonomy of entities can be represented as a semantic net in a form of a tree (systemic grammars) [15]. This type of semantic net is useful in the building of semantic structures (see section 5.1.1).

2.3 Issues

2.3.1 Usefulness of Information

There is one issue that should be addressed before any work is done on the creation of a system. How useful is the information gained from the the free-format parts of the message? It is the opinion of this author and others [9] that the free-format part does not significantly contribute new information in a large percentage of the messages. For most of the significant information there exists a fixed-format line in which the information can be encoded. The free-format text comes into play when either a novice or expert message-sender does not use the fixed format through ignorance or laziness; it is easier to use narrative in describing an event than to remember the message formats. In the short run, it might be better to concentrate on making the message-senders learn to do the formatting better. However, in the long run, NLP should become a factor for easing the rigidity of partly formatted messages.

2.3.2 Technology Transfer

Because AI and, in particular, NLP are at the beginning stages of development, it is difficult to "transfer the technology." In fact, there is very little technology transfer within the field. Almost all the AI systems are prototypes only. The vast majority of the systems have been written from scratch, and the many recurring problems in the field are solved individually each time. At this point NLP has only some guiding principles:

- The ability of the system is directly proportional to the amount of precanned knowledge in the system.

- Semantic analysis is the key to understanding language, but knowledge of syntax is necessary and useful.

- Expectation-based parsing is to date the most successful method.

- Many types of knowledge and knowledge structures are needed in understanding.

2.3.3 Expert Systems

All successful AI programs have had one thing in common. The program was built by a computer expert, but with active help of an expert in the field the program addressed. For instance, MYCIN, a program to diagnose blood diseases, was written by a computer science graduate student in conjunction with a medical doctor specializing in blood diseases. It is highly recommended that, for the project of understanding Navy messages, a resident expert on naval procedures be readily available for consulting. If this expertise is not provided, then the project is probably doomed to failure. If the programmer does not understand the events in the messages, then his program cannot possibly understand them either.

3. LANGUAGE ASPECTS

3.1 Introduction

This section will discuss each syntactic category relative to understanding NTM. There are eight parts of speech: verbs, nouns, adjectives, adverbs, prepositions, conjunctions, pronouns, and interjections. All categories are used in NTM except interjections. The discussion will outline the importance, the frequency of occurrence, and the role of the category in the understanding process.

The understanding process has the ultimate goal of combining the structures that are suggested by the words in the sentences. Parsing helps that goal by giving some clues to which structures should attach to which other structures. Nouns fill the slots of the verb structures; adjectives modify noun structures; adverbs modify verb structures. Prepositions direct where noun structures fit in the verb structures. An incorrect parse results in one piece filling or modifying in the wrong structure.

3.2 Verbs and verbals

It has been noted that the verb is the most important part of speech in understanding. In understanding the NTM, there is an advantage in that the vast majority of the sentences have the main verb expressed in past tense. This helps the understanding process, because it provides a strong heuristic for choosing the verb.

Although the verb might be easy to choose, there is a slight complication. Some sentences are in the passive voice. Normally, the passive voice is easily recognized by the existence of the auxiliary verb. Unfortunately, the sentences frequently have the auxiliary omitted, as in sentences 1, 2, and 5 in example 1.3. The omission creates a slight syntactic similarity between the normal and passive voice.

Contact submerged. -> The contact submerged.
Contact gained. -> The contact was gained by us.

Luckily, the passive can be detected easily by a sentence having a transitive verb with only one noun phrase being before the verb. If the verb can be either intransitive or transitive, then it is assumed the verb is intransitive and, therefore, not passive.

The second most common tense is the past progressive, which is similar to the passive past tense in that the auxiliary is omitted. Both omissions make possible a confusion between the main verb and a participle serving as an adjective. (Gerunds are rare in NTM.) For instance:

Deployed modified delta pattern.

The solution to resolving the confusion is a combination between syntactic heuristics and semantic analysis. More detail will be provided in sections 3.3 and 5.1.1.

The present tense occurs rarely. It is a tense to be picked last in the choice for main verb. Frequently this situation occurs when the sentence has "intend" as the main verb. Future and perfect tenses (i.e., future, future perfect, past perfect, etc.) are easily recognized syntactically. The auxiliary or modal must be physically present in the text. These tenses are used rarely.

Some sentences do not have a verb. This is the result of a copula verb being omitted in the sentence. Sentence 3 is an example. The forms "there was" and "it was" are the texts typically omitted. It can be assumed that all sentences without a verb are of the copula form.

The infinitive form of the verb does occur frequently. The *infinitive form does occur* without the "to," but only rarely, as in the sentence:

Intend make sweep. -> We intend to make a sweep.

The infinitive form is used normally as an adverb to the main verb. This gives a good clue that the main verb is just before the "to" (backward prediction). Infinitives serving as noun phrases were not found in the sample NTM; they would contradict the backward prediction.

3.3 Noun groups, Adjectives, and Nouns

One of the most difficult parts of parsing arises in deciding the boundaries of a noun group. This problem arises because each word normally can serve in several syntactic roles. In this series of sentences [7],

```
The US forces fight.
The US forces fight in Vietnam.
The US forces fight in Vietnam is costly.
```

each has FORCES and FIGHT in different syntactic roles. The first sentence has FORCES as a verb and FIGHT as a noun. Sentence three has both FORCES and FIGHT as nouns.

The NTM have several aspects to them that hinder noun phrase recognition. On the other hand, the NTM have a certain style and content which help the process.

The noun phrase is composed of a set of adjectives, adverbs, and nouns. The noun can also be modified by an infinitive or a prepositional phrase. For parsing, there are several basic heuristics of syntax.

- adverbs precede most adjectives.
- articles precede numbers.
- numbers precede the other adjectives.
- adjectives precede nouns.

The adjective is tightly bound to the processing of the noun phrase. In normal prose, articles and numbers give a good clue to where a noun phrase starts, because they are unambiguously adjectives (they don't belong to other parts of speech), and they are the first word in the noun phrase. Unfortunately, articles are used rarely in the NTM. This is especially unfortunate for sentences that have no verb, as in sentence 3. All that is left is a bag of noun phrases, and there is little to separate them syntactically.

It has been said that the participle form can be used as an adjective. To distinguish it from the other roles that the participle can play, the following heuristics can be used.

1. If the participle is preceded (immediately) by an adjective, then the participle is an adjective.

2. If it precedes an adverb, then it is probable that it is a verb.

3. If it is preceded by a noun, then it is a verb.

The difficulty arises in the following situations. Notice the heuristics do not work, because the articles are omitted.

Finished dropping last barrier.
Saw running lights.

Semantic analysis is needed in these situations.

In normal prose, the infinitive and participle phrase can serve also as an adjective. However, this situation does not occur in the NTM, and it should be noted that they do not occur as noun phrases either.

One helpful aspect of the NTM is the restricted use of verb tenses. Since the present tense is rarely used, a word in the present tense form that can be either a noun or verb is likely to be a noun. For instance, this heuristic would give added evidence in example 1.3 that FLARES and ATTACKS were nouns.

One situation that does arise occasionally will cause some difficulty. Omitting "of" from a noun phrase does occur, e.g., REMAINDER ONSTA TIME. Using the heuristics above, the likely result will be parsing of two noun phrases. The resolution will come only if the semantics dictate that there are too many objects, and they do not fit in.

The style of the NTM gives other weak heuristics choosing the syntactic category: 1) If two participles are together, then the first participle is the main verb; 2) A word in the infinitive form without a preceding "to" is part of the noun phrase. In general, the noun phrase is short: the average number of words is about two.

3.4 Adverbs

The adverb has essentially two forms: the single word modifier and the infinitive phrase. The single word modifier is used occasionally. These modifiers are not of major importance, and they are relatively easy to incorporate in the semantic structures. In the NTM, the adverbs are just before the verb, and the verb is typically at the beginning of the sentence. The adverb can also modify an adjective. In this case the adverb does not serve a different role than the adjective and can be processed in the same manner (i.e., they modify the noun structure).

The adverbial infinitive phrase is simple to detect syntactically. (This is unless the "to" is deleted.) The more difficult part is the merging of the conceptual structures of both the main verb and infinitive verb. The nesting of the verbs can be several levels. For example:

Continued to investigate to insure

More detail on the semantics will be discussed later.

3.5 Prepositions

The use of prepositions in the NTM does not vary significantly from the use in normal prose. The preposition plays primarily a case marking role. The meaning of the preposition is determined by the verb, as in the following sentences.

John cheated on Mary.
John landed on Mary.

However, one can consider each preposition with a default meaning. Each preposition has a standard definition. For example, "on" is defined as physically on top-of; "in" is defined to be contained-in; "with" is defined as "using" for an inanimate object and "accompanying" for an animate object (platforms are considered animate). For each verb there corresponds a script, a situation, or a primitive act; and they have slots to fill. The prepositions give some indication what slots to fill. Take the situation $ATTACK. Its slots are: ATTACKER, VICTIM, WEAPON, ACCOMPLICE, LOCATION OF VICTIM, LOCATION OF ATTACKER, DIRECTION

OF ATTACK, REASON OF ATTACK, TIME OF ATTACK. The situation will suggest the possible slots that a particular object of the preposition will fill. E.g., the object of "on" will fill the victim slot if the object is a platform. The situation does not suggest anything for "with," because the preposition's default definition fills in the correct slots.

One slight difficulty arises with verbs that include a preposition as part of their definition, e.g., "check in" and "follow up." Luckily, these verbs occur rarely; and when they do occur, two prepositions are in juxtaposition, which signals this type of verb. For example:

Checked in with ax.

The other difficulty arises with a preposition occurring twice in the same sentence. For instance:

Visual contact on sub on surface.

This is resolved by the properties of objects. In this example the situation suggested is $CONTACT, which has the slots: CONTACTER, CONTACTEE, SENSOR, LOCATION OF CONTACTER, LOCATION OF CONTACTEE, LOCATION OF SENSOR, TIME. The objects naturally fill the correct slots: SUB is a platform and, therefore, the CONTACTEE; "on" normally suggests a location so SURFACE is the location of CONTACTEE.

3.6 Conjunctions

The conjunction is classified as either coordinate or subordinate. Coordinating conjunctions are words that join two equal forms; subordinate conjunctions join a dependent clause to the main clause. The NTM use almost exclusively the coordinating conjunction "and." Any other coordinating conjunction is rare. (Only two instances of another conjunction were found in the sample.) The subordinate conjunctions in the NTM are primarily adverbs of time, e.g., "while," "after," and "before."

The treatment of conjunctions involves special processing procedures. The processing operates in terms of levels. The text following a conjunction varies from a noun phrase to a complete sentence. When the conjunction is encountered, the processing is saved, and a conceptually or syntactically complete unit is expected. As soon as that unit is processed, the previous processing is restored, then the structure produced at the sub-level is merged with the structure at the higher-level.

Let us take a specific example to further specify the process.

Proceeded to area and sighted periscope.

"Proceeded to area" is converted to the structure:

(PTRANS (ACTOR ms) (OBJECT ms) (TO (LOCATION (AREA areal))))

Since there already exists a verb when "and" is encountered, the expectation is another action with the message-sender as the actor. The parsing proceeds at the sub-level. This means that a new structure is built and not incorporated in the existing structure. "Sighted" produces a verb structure and "periscope" is put in that structure. The sentence ends, making that sub-level terminate. The sub-level returns a verb structure which cannot be incorporated into the previous structure; thus, the implied subject is copied into the second structure.

Most of the time the use of conjunctions in the NTM is syntactically correct. However, there are enough instances of improper use of "and" to warrant some attention being paid to methods of checking the reasonableness of the interpretation. The problem lies in the implied use of the subject. That is, compound sentences use a common subject for the two verbs, as in the example. The verbs "proceed" and "sight" have the same subject. However, as we saw in example 2.1, this syntactic rule can be violated.

> Surface contact turned away upon approach and began stern chase.

How could the surface contact (hostile ship) turn away and chase at the same time? It can't. The subject of "began stern chase" is the platform who sent the message and not the hostile ship. Only with the understanding that a platform cannot do an $EVADE with a $PURSUE, can the normal syntactic heuristic be overruled and the correct interpretation taken.

The next example has a compound object.

Tass ship had range and lob which was not passed.

When the "and" is encountered the semantic structure is:

```
(KNOW (ACTOR (PLATFORM (INSTANCE (SHIP shipl))
                       (SENSORS (INSTANCE TASS))))
      (FACT (LOCATION (PART platl) (DIST SPECIFIC))))
```

The same situation exists as in the other example except a noun phrase with a relative clause is after the "and." It cannot stand alone; it must be incorporated in the previous structure. Since LOB is similar to RANGE, the two can be combined to be the object of the KNOW, and the relative clause can predicate both.

3.7 Pronouns

The NTM do not use pronouns in any significant manner. The primary use is syntactic. The syntactic constructs "this is," "there are," and "it is" are some places they occur. These constructs are also not used very much; usually the sentences that have no verb are sentences with the constructs, but the construct is physically omitted. For instance:

There is not joy on buoy pattern.
-> No joy on buoy pattern.

The syntactic constructs have no meaning and, therefore, their omission does not present a problem. Relative pronouns (which, who, that) are the other pronouns used in NTM. They are also simply dealt with.

Nevertheless, a major problem in NLP is pronoun resolution. Although the NTM do not have pronouns, the problem does not go away, but it shifts to another problem. The problem becomes the resolution of implied objects of verbs. For example, the sentence "Attacked with asroc" omits both the subject and the direct object. Normally in prose, if the subject and object had been mentioned before, then they would be referred to by pronouns. The NTM take the process one step further; they omit the reference completely. (The Japanese language does the same thing.) Nevertheless, to understand the sentence, some decision must be made in resolving the subject and direct object. Although there are no pronouns, the problem is still basically pronoun resolution.

3.8 Syntactic Constructs

Something must be briefly said of the range of syntactic diversity in the NTM. The messages have a definite style. Not all constructs possible in English are used. More specifically, all sentences are in the declarative form: there are no sentences that are imperative, interrogative, or exclamatory. In addition, quotations, absolute phrases, and parenthetical expressions are not used. Constructs such as appositives, unrestricted relative clauses, gerunds, and noun clauses occur rarely. Introductory phrases and relative clauses are occasionally used. It should be noted that the introductory phrase is usually in the form of "at <time>." Compound sentences are used, but the vast majority of the sentences are of the simple form. Complex sentences are used occasionally, and, accordingly, there are a small number of compound-complex sentences in the messages.

4. SPECIAL ISSUES

4.1 Missing Words

It is the opinion of the author that a major problem in understanding Navy tactical messages is the overwhelming amount of special abbreviations and acronyms that are used. People unfamiliar (notably the author) with the Navy jargon despite an understanding of the world and the English language have a great difficulty understanding sentences with jargon words. One major weakness of NLP has been the use of small vocabularies. No NLP system has had a significant number of words known to it (over a couple hundred words). For a practical system, there must be a significant amount of the abbreviations and acronyms known to the system.

It is unreasonable to assume that all the words in a message will be in the system's lexicon. The number of acronyms in the Navy can easily fill several books. Even restricting the domain will not solve the problem.

Some number of words are names of ships, countries, weapons, and devices. The names and class names are easily over a thousand. Call names also introduce unknown words into the messages. It is unreasonable for the system to know these names.

The most troublesome aspect of the NTM is the use of abbreviations for common words. E.g., vic = vicinity, rtn = return, unident = unidentified, etc. Some of the common abbreviations (semi-Navy terms) can be put into the system, e.g., sub = submarine, tgt = target, flt = flight, ctc = contact, pos = position, etc. But it never can be assumed that all abbreviations can be known; any message-sender can arbitrarily abbreviate any word (and they do) to shorten the message length. People are good at guessing the possible words; whereas the computer guessing at the word would become computationally expensive and probably incorrect. A large percentage of the abbreviations are just the first part of the word, e.g., on sta = on station, approx = approximate, etc. These words might be found through a hashed searched, but this operation is also expensive.

Misspelling of words occurs in all human written communication. The NTM are no exception. Examples: persecuting = prosecuting, form = from, fo = of, etc. There are known algorithms for spelling correction [12]. However, since there are several sources of unknown words: abbreviations, acronyms, and names, it probably is fruitless or unwise to try to use those techniques. Some experimenting should be done on the preliminary systems to actually determine if it is useful. (It is simple enough to implement.)

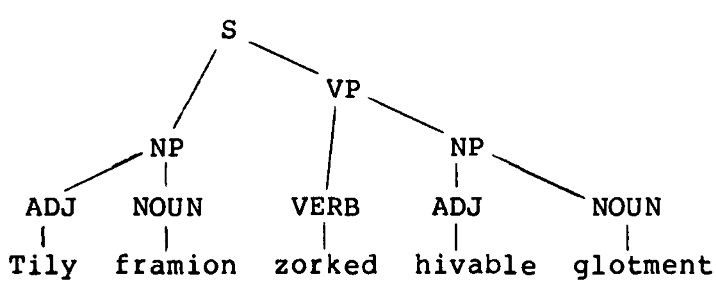
4.1.1 Using Expectations

One approach to handling missing words is along the lines of FOUL-UP [8, 4]. The basic idea is to use the expectations that are generated from the sentence to classify the word. For example, "visual exchange of call signs indicate USS framus."

Although "framus" is not known as a word, we hypothesize that it is a ship's name. USS is an adjective, which gives a high confidence expectation that the next word is a ship name. The semantics of what can exchange call signs adds to the confidence.

Semantics of the surrounding context can also guide the further search to lexical analysis of an unknown word. Sometimes the words get run together: 5IN54 GUNS. A procedure could try to break apart the word into smaller lexical items. The process is computationally expensive, but the search can be guided because the context restricts the possibilities.

Besides semantic expectations, knowledge about morphology can be used to try to guess a word's meaning. In English, the primary morphological clues come from *suffixes*. Endings such as "ing," "ed," "ment," give a definite clue as to what syntactic category the word is in. The following is an example of a nonsense sentence with morphological endings.



Just with the morphological data the sentence can be parsed (but not understood). However, some endings can serve two syntactic categories, e.g., -ic (noun, adjective). These endings still can rule out possibilities. Specifically, -ic can rule out the word being a verb.

It should be noted in passing, morphological analysis can save storing in the lexicon the variant forms of the word. A simple procedure to strip endings can be devised.

The NTM have certain features which should be noted as morphological. Many abbreviations do not contain vowels, e.g., SLCM, WWMCSS, LST. Any word containing a digit connotes an object, a class, or an identification name: CV-15, APS-165. Time is represented always by four or six digits, a Z, checksum digit. There are numbers which connote possible meanings: degrees are from 0 to 360; speed is from 1 to 35; a number in the thousands might be yards. All these are heuristics that give some clue to what something might mean.

4.1.2 Confidences

Another thing can be done in dealing with the unknown words in conjunction with the expectation based approach. Since the expectations are only heuristics, which are subject to errors, the idea is to associate every "guess" with a confidence factor. Take the text:

XYZ 404 was on photo mission

The following rule could be applied. If 1) the unknown word is followed by a number, 2) that number is not associated with some other word, 3) there is a need for a platform then, with confidence .9, the word is a call sign. The result would be that XYZ 404 would be assumed to be a platform with confidence .9 (confidence -1 to +1).

The interpretation of text can be given a reliability weighting either for individual facts or the text as a whole. Matching of knowledge structures would then have a margin of error. If the confidence was below a certain threshold the sentence or the entire message would not be considered understood.

4.2 Interfacing with Reasoning Systems

The ideal situation for integration between a NLP system and other reasoning systems is for them to be combined into the same program. Otherwise, the representation constructed by the NLP system must be put in a form so that another system can retrieve it. This is not difficult.

Some current prototype systems represent information in the form "aRb" (R is a relation on objects a and b). The NLP representation proposed is more like a labeled n-tuple. The difference is artificial. One solution is to convert the NLP representation to the simpler form. This is not recommended because it makes the information more voluminous. The other solution is to have the reasoning programs use the same formalism, or at least be able to extract the wanted information. It is recommended that all future reasoning systems use the NLP formalism because it is a more flexible representation.

Nevertheless, this issue is not seen as critical for prototype systems.

5. NLP APPLIED TO NAVAL TACTICAL MESSAGES

5.1 Building Semantic Structures

The understanding process can be seen as three basic processes: building, filling, and matching. As parsing proceeds, each word contributes a semantic structure, which serves as a building block for construction of the total structure or the "meaning" of the message. The understanding process takes the building blocks and tries to fit them together to make a coherent structure.

One major process, filling, directly connects two structures together. Filling always involves putting a subordinate structure into a suprastructure. More specifically, adjective structures are put in noun structures, and noun structures are put in verb structures. The building process is collecting the structures and supervising the filling process. The filling process involves checking all unfilled slots in existing structures and choosing the correct structures to fill those slots.

5.1.1 Noun Phrases

Take the noun group, "surface combatant." "Surface" is defined as three possible semantic structures. "Surface" can be the adjective structure (INSTANCE SHIP), the noun structure (LOCATION (PART SURFACE)), or the verb structure (PTRANS (TO SURFACE) (FROM (LOCATION (INSIDE OCEAN)))) In other words, surface has three meanings: a type of platform, a location, and a movement from one location to another. The word "combatant" is the noun structure (PLATFORM (SIDE HOSTILE)). Since platform has other slots:

(speed, propel-type, category, side, etc.), the filling process will try to fill them. Since "surface" serves as an adjective in this instance, (INSTANCE SHIP) is chosen. It matches the slot of the platform, and the resulting structure is:

(PLATFORM (INSTANCE SHIP) (SIDE HOSTILE))

In a different example, "Sonobuoy contact," the filling process becomes slightly more complex. The only representation for "sonobuoy" is the noun structure BUOY. "Contact" has three senses: two noun structures PLATFORM and $CONTACT, and the verb structure $COMMUNICATION. There are two problems: there is no direct match of BUOY in any of the slots of the structures and there must be a choice between the structures.

The first heuristic is to try to match the noun structures. Since there is no direct match, the semantic network must be employed. Traveling up the hierarchy, one finds BUOY is an INSTANCE of SONAR and SONAR is an INSTANCE of SENSOR. Both PLATFORM and $CONTACT have the slot SENSOR. However, the platform-slot SENSOR represents a part of a platform (i.e., platforms have sensors), whereas the $CONTACT-slot SENSOR represents an attribute. Using this distinction, the resulting structure is:

($CONTACT (SENSOR (SONAR (INSTANCE BUOY))))

5.1.2 Verb Phrases

Verb phrases can contain one or two infinitives or participles that are adverbial modifiers. Semantically, those phrases represent statements about goals or a further refinement on the time of the verbs.

Goal:
  Continued to investigate to insure
  Remaining to attempt to regain

Time:
  Finished dropping
  Began chasing
  Commenced snorkeling
  Continued to orbit

The building of the semantic structures for these phrases involves determining a dominate verb, which may not be the main verb. The verb can be classified into several broad categories: time, goal, general, and action.

time: continue, finish, begin, commence
goal: insure, attempt, tasked, directed, intend
general: establish, conduct, revert, be, do, appear
action: drop, chase, snorkel, orbit, remain

The dominance order of the types is: action, general, goal, and time. If there is an action verb in the phrase, then the semantic structure is built around that verb. General verbs usually derive their meaning from an action verb which is in adverb form or a noun phrase.

Appeared to be tracking
Conducted an attack

Goal verbs set up the goal primitives, which have two slots for verb-like structures. If there is an accompanying action verb then it goes into the ACTION slot of the primitive, and the goal verb goes into the INTENTION slot. Time verbs can serve to modify the other dominating verbs.

5.2 Expectation-based Parsing

For a long time in NLP, a major problem of parsing was ambiguity. With the advent of expectation-based parsing, ambiguity has been relegated to a minor role. Another larger problem, search, i.e., the looking for relevant information, has been helped by the approach also.

Briefly, expectation-based parsing is a process where the structures attached to each word suggest what to expect. The expectation can be forward or backward. Backward expectation predicts something previously encountered (to the left), and forward predicts something that has not been parsed (to the right). There are two types of expectations: predicted and promised. (In reality there is a continuum between the two.) The predicted expectation does not have to occur, whereas the promised expectation must occur or something is considered wrong.

For example, a simple expectation of a small number (< 5) predicts a plural noun, as in example 1.3. "Bearing" predicts the word "degrees." At the situation level, the situation $DETECT predicts $SEARCH; $SEARCH predicts $CONTACT; $CONTACT predicts $IDENTIFY.

"Bearing" also has a promised expectation. Almost always bearing represents an angle between 0 and 360. It can be said that if one chooses that interpretation, then it promises a number or direction. However, it can be seen, as the example shows, to be interpreted as something else.

Fired 2 harpoon brg/range mode.

The example uses bearing (brg) with range to form an adjective for mode. The missing of the promised expectation signals a misinterpretation of the text. Other words have similar expectations: miles, speed, range, etc.

The required slots in a situation are promised expectations when the situation is trying to be instantiated. When interpreting a sentence as part of a situation, the absence of an important part signals that the situation is not correct. The following is an example of a false lead in expectations.

Flares sighted. Gained attack criteria.

The first sentence invokes the situation $DETECT, which then predicts $CONTACT. One of the definitions of gain is consistent with $CONTACT as in the situation "Gained contact," however, the rest of the sentence does not confirm the situation. Because the situation must be promised a platform to be invoked, the situation will fail, and other possibilities must be tried.

Sometimes the promised expectation is backward, that is, the word predicts something before it. If the backward promised expectation is not satisfied immediately, then the sentence must be reparsed.

As the structures are built, the decisions made in construction have a confidence associated with them. Confirming of expectations adds to the confidence weight. If a reinterpretation is needed, then the decision with the lowest confidence (relevant to the interpretation) is reversed.

5.3 Sample Knowledge Structures

5.3.1 Primitives

5.3.1.1 Space

Spatial Domain: representing the spatial relationships between objects.

Primitive Acts

```
PTRANS = Physical Transfer (change in location)
  cases:
   ACTOR - The person doing the action
   TO - The location changed to
   FROM - The location changed from
   OBJECT - object transferred by actor

ROTATE = Rotation (change in orientation)
  cases:
   ACTOR - The person doing the action
   TO - The direction the object is facing after Rotate
   FROM - The direction before Rotate
```

Primitive States

```
ORIENTATION
  cases:
   PART - the object whose orientation is being described
   DIRECTION - the direction the object is facing

DISTANCE
  cases:
   AMOUNT - the number of units
   UNITS - the type of units

DIRECTION
  cases:
   BEAR - bearing
   REFERENCE - the reference location
   TO - the location used as a direction
          (i.e., toward the mountains)
```

```
LOCATION
  cases:
   PART - The object whose location is being described
   DISTANCE - the distance from a location
   LOCATION - reference location
   DIRECTION - the direction of the object with
               respect to the reference
   INSIDE - a location somewhere inside the object
   OUTSIDE - outside object
   SURFACE - on the surface
```

Sample Encodings

Closed Kashin to 21k yards.

```
(PTRANS (ACTOR MS) (TO LOC1) (FROM SPECIFIC))
(LOCATION (PART LOC1)
          (DISTANCE (AMOUNT 21000) (UNITS YARDS))
          (LOCATION (PART KASHIN)))
```

Proceeded to area.

```
(PTRANS (ACTOR MS) (TO (AREA AREA1)) (FROM SPECIFIC))
```

Contact turned away.

```
(ROTATE (ACTOR CONTACT1) (DIRECTION (BEAR 0) (REFERENCE O1)))
(ORIENTATION (PART O1) (DIRECTION (TO CONTACT1)
             (REFERENCE MS))
```

Sample inferences

```
If (PTRANS (ACTOR x) (TO x1))
 then (LOCATION (PART x) (LOCATION x1))
```

I.e., If something moves then its location changes.

If a platform ROTATES, then there is a change of course

5.3.1.2 Motion

Motion Domain: representing the general movement of objects.

Primitive Acts

```
MOVE = The movement of an object in some manner
  cases:
    AREA - area which movement occurred in
    TO - the location which movement was toward
    FROM - the location which movement was away from
    DIRECTION -direction in which the movement occurred
    SPEED - the velocity of the movement
    ACTOR - the person who did the movement
```

Sample Encodings

Investigated area

```
(WATCH (ACTOR ms) (OBJECT SPECIFIC))
(MOVE (ACTOR ms) (AREA areal))
```

Tracking South speed 17.

```
(ATTEND (ACTOR ms) (OBJECT contactl))
(MOVE (ACTOR contactl) (DIRECTION (BEAR 180) (REFERENCE EARTH))
      (SPEED (AMOUNT 17) (UNITS KNOTS)))
```

Sample Inferences

```
If (MOVE (ACTOR x) (DIRECTION dl (REFERENCE EARTH))
 then (ORIENTATION (PART x)
                   (DIRECTION dl (REFERENCE EARTH)))
```

I.e., if x is moving in a direction, then x is pointed that way. E.g., Ship A is moving south, the starboard side is west.

5.3.1.3 Goal

Goal Domain: representing motivations and goals of platforms.

<u>Primitive Acts</u>

```
PURSUE = Pursuing a goal actively
  cases:
    ACTOR - platform pursuing the goal
    ACTION - action that will enable goal success
    INTENTION - the goal trying to be achieved

ABANDON = Abandon a goal
  cases:
    ACTOR - platform abandoning the goal
    GOAL - the abandoned goal
```

<u>Primitive States</u>

```
GOALSTATUS = the success, pursuit, or failure
  cases:
    GOAL - goal
    STATUS - state of goal
```

<u>Sample encodings</u>

Remaining in area to attempt to regain contact.

```
(PURSUE (ACTOR ms) (ACTION (MOVE (ACTOR ms) (AREA areal)))
        (INTENTION ($CONTACT (CONTACTER ms)
                             (CONTACTEE specific))))
```

Attempted active pattern with negative results.

```
(PURSUE (ACTOR ms)
        (ACTION ($SEARCH (SEARCHER ms)
                         (SENSOR (SENSOR (TYPE ACTIVE)))
        (INTENTION ($CONTACT (CONTACTER ms))
                   (NAME goall)))
(GOALSTATUS (GOAL goall) (STATUS FAIL))
```

<u>Sample Inferences</u>

```
If (GOALSTATUS (GOAL gl) (STATUS FAIL)) and
   (PURSUE (INTENTION gl))
 then either (PURSUE (ACTOR x) (ACTION al) (INTENTION gl))
      or (ABANDON (ACTOR x) (GOAL gl))
```

I.e., If a goal fails, then either try something else or abandon the goal.

5.3.1.4 Vision

Vision Domain: representing perception

Primitive Acts

```
ATTEND = the perceiving of an object.
  cases:
   ACTOR - the person perceiving
   OBJECT - the object perceived
   SENSOR - the sense being used
     e.g., visual, radar, sonar, MAD
   DIRECTION - the direction of perceiving

WATCH = the active looking for something
  cases:
   ACTOR - the person watching
   OBJECT - the object being looked for
   SENSOR - the sensor being used to watch
```

Sample encodings

Sighted flares.

(ATTEND (ACTOR ms) (OBJECT flares) (SENSOR VISUAL))

Positive MAD contact held by helo.

```
(ATTEND (ACTOR helocopter1) (OBJECT contact1)
        (SENSOR MAD))
```

Investigated area.

```
(MOVE (ACTOR ms) (AREA area1)
(WATCH (ACTOR ms) (OBJECT specific))
```

Sample inferences

```
If (ATTEND (ACTOR x1) (OBJECT o1) (LOCATION l1)
           (TIME t1))
  then (LOCATION (PART o1)
                 (LOCATION l1) (TIME t1))
```

I.e., if an object is seen in a location then the location of the object is there.

If a part of an object is seen, then the object has been seen.
E.g., if a periscope is seen, then the sub has been seen.

If sub is submerged, it cannot be seen visually.

5.3.1.5 Resource

Resource Domain: representing the use and disposition of resources.

Primitive Acts

CONSUME = the consumption of resources
USE = the use of recoverable resources
REPLENISH = the renewal of resources

Sample Encodings

While refueling alongside USS Bronstein.

```
(CAUSE (PTRANS (ACTOR ms) (OBJECT fuel) (TO ms) (FROM shipl))
       (REPLENISH (ACTOR ms) (RESOURCE fuel)))
```

40 rounds expended.

```
(CONSUME (ACTOR ms)
       (RESOURCE (WEAPON (INSTANCE BULLETS)
                         (AMOUNT 40) (UNITS ROUNDS))))
```

Deployed sonobuoys.

```
(USE (ACTOR ms) (RESOURCE (SENSOR (INSTANCE BUOY))))
```

Sample inferences

If a resource is CONSUMEd, then the resource cannot be used again.

If a resource is REPLENISHed, then the resource can be used.

5.3.1.6 Knowledge

Knowledge Domain: representing the knowledge known or gained by platforms.

Primitive Acts

MTRANS = transfer of knowledge
DECIDE = a decision being made

Primitive States

KNOW = the state of knowing a fact

Example Encodings

Visually identified as J. P. Jones.

```
(DECIDE (ACTOR ms)
        (DECISION (KNOW (ACTOR ms)
                        (FACT (PLATFORM (INSTANCE SHIP)
                                        (NAME JPJONES))))))
```

Contact failed to acknowledge challenge.

```
(MTRANS (ACTOR contactl) (MANNER NEG)
        (KNOWLEDGE (KNOW (ACTOR contactl)
                         (FACT ($CHALLENGE (CHALLENGER ms)
                                           (CHALLENGEE contactl)))))
```

5.3.2 Situations

AVOIDDETECTION

Description

Taking some action or not an action to result in another platform not knowing where it is.

Examples

Turned off running lights
Radar emcon
Sub went sinker

Representation

```
$AVOIDDETECTION(detecter,detectee)
  (PURSUE (ACTION a1)
          (INTENTION (KNOW (MANNER NEG) (ACTOR detector)
                           (FACT (LOCATION (PART detectee))))))

 a1 = (USE (ACTOR detectee) (MANNER NEG)
          (RESOURCE (SENSOR (TYPE ACTIVE))))
   or (PTRANS (ACTOR detectee) (TO (LOCATION (INSIDE OCEAN))))

Inferences:
    detectee is hostile.
Predictions:
   Forward - $SEARCH, $CONTACT
   Backward - $CONTACT
```

SEARCH

Description

The searching of an area to see what is there.

Examples

Investigated area
Mad search
Made sweep of area

Representation

```
$SEARCH(searcher,area,sensor)

(WATCH (ACTOR searcher) (SENSOR sensor))
(MOVE (ACTOR searcher) (AREA area))
```

Inferences:
If a platform is detectable by sensor and it is in the area, then it is probable that the searcher will see it.

Predictions:
Forward - $CONTACT, $LAYBUOY
Backward - $DETECT

LAYBUOY

Description

The laying or dropping of buoys for the purpose of detecting or fixing a position of a submarine.

Examples

Deployed modified delta pattern
Laid 10 buoy
Sonobuoy drops

Representation

```
$LAYBUOY(deployer,buoy,loc)

(PURSUE (ACTOR ms)
        (ACTION (CAUSE (DO (ACTOR ms))
                    (PTRANS (OBJECT buoy)
                            (TO (LOCATION (PART loc))
                                  (LOCATION (AREA SURFACE))))))
        (INTENTION ($CONTACT (CONTACTER ms)
```

```
                              (CONTACTEE sub)
                              (LOCATION (PART loc)))))
(USE (ACTOR ms) (RESOURCE buoy))
```

Inferences:
   If there is no contact, then the sub is not in the area

Predictions:
   Forward - $CONTACT, COMMRELAY
   Backward - $CONTACT, $DETECT, $AVOIDDETECTION

CONTACT

Description

The locating of a platform

Examples

Contact gained
USS Smith sighted on surface
Positive MAD contact

Representation

```
$CONTACT(contacter,contactee,sensor)

(ATTEND (ACTOR ms) (OBJECT contactee) (SENSOR sensor)
        (LOCATION (PART contactee)))
```

Predictions:
   Forward - $IDENTIFY, $ATTACK
   Backward - $SEARCH, $LAYBUOY

5.3.3 Scripts

```
COMMUNICATION-RELAY(relayer,receiver,buoy)

 1. ($LAYBUOY (ACTOR relayer) (BUOY buoy))

 2. (MOVE (ACTOR relayer) (LOCATION (AREA area)))
    ($COMMUNICATION (SENDER buoy) (RECEIVER relayer)
                    (INFORMATION info))
    ($COMMUNICATION (SENDER relayer) (RECEIVER receiver)
                    (INFORMATION info))
    (LOCATION (CENTER area) (LOCATION (PART buoy)))


PATROL(patroller,command,area)

1. ($COMMAND (COMMANDER command) (COMMANDEE patroller)
             (ACTION ($PATROL1 (PATROLLER patroller)
                                (AREA area))))

PATROL1
2. (PTRANS (ACTOR patroller) (TO (AREA area)) (FROM base))

3. ($SEARCH (SEARCHER patroller) (AREA area))

4. (PTRANS (ACTOR patroller) (TO base))

5. ($COMMUNICATE (SENDER patroller) (RECEIVER command))
```

5.3.4 Example 1.3

The following is a partial representation of the example, after the understanding process is done.

```
(ATTEND (ACTOR ms) (OBJECT (PHYSOBJ (INSTANCE flare) (AMOUNT 2)))
         (SENSOR VISUAL)
        (DIRECTION (BEAR 180) (ACCUR COMP)))

(LOCATION (PART (AREA areal))
          (DIRECTION (BEAR 226) (ACCUR DEGREES))
          (DISTANCE (AMOUNT 5) (UNITS MILES))
          (LOCATION (PART (SHIP (NAME KITTYHAWK)))))

(AREA (NAME areal) (RADIUS (AMOUNT 5) (UNITS MILES)))

($DETECT (DETECTER ms) (DETECTEE contactl) (TIME 230704))

($SEARCH (SEARCHER ms) (AREA areal)
         (TIME (BEGIN 230704) (END 230735)))

($CONTACT (CONTACTER ms) (CONTACTEE contactl)
          (TIME 230735))

($IDENTIFY (IDENTIFIER ms) (IDENTIFEE contactl)
           (IDENTIFICATION SUBMARINE)
           (CONFIDENCE 3))

($TRACK (TRACKER ms) (TRACKEE contactl)
        (DIRECTION (BEAR 135) (ACCUR COMP))
        (SPEED (AMOUNT 17) (UNITS KNOTS)))

($ATTACK (ATTACKER ms) (ATTACKEE contactl)
         (WEAPON (INSTANCE (MISSILE (INSTANCE ASROC)))
                 (AMOUNT 2)))

($LOSTCONTACT (CONTACTER ms) (CONTACTEE contactl) (TIME 230743))
```

REFERENCES

[1]
R. Bechtel, P. Morris, and D. Kibler.
STAMMER2: A Production System for Tactical Situation Assessment.
Technical Report, SDC Integrated Services, August, 1979.
(Also in progress as NOSC Technical Document 298.).

[2]
R. J. Brachman.
A Structural Paradigm for Representing Knowledge.
PhD thesis, Harvard University, 1978.
BBN Report 3805.

[3]
Richard Burton.
Semantic Grammars.
PhD thesis, University of California, Irvine, 1976.

[4]
Jaime G. Carbonell.
Towards a Self-Extending Parser.
In 17th Annual Meeting of the Association for Computational Lingustics, pages 3-7. Association for Computational Linguistics, University of California, San Diego, August, 1979.

[5]
Robin A. Dillard.
New Methodologies for Automated Data Fusion Processing.
Technical Report 364, Naval Ocean Systems Center, September, 1978.

[6]
C. Fillmore.
The Case for Case.
In Bach and Harms, editors, Universals in Linguistics Theory, pages 1-90. Holt, Rinehart, and Winston, 1968.

[7]
Anatole Gershman.
Conceptual Analysis of Noun Groups in English.
In 5th International Joint Conference of Artificial Intelligence, pages 132-138. MIT, 1978.

[8]
Richard H. Granger Jr.
FOUL-UP: A Program that figures out meanings of words from context.
In _5th International Conference of Artificial Intelligence_, pages 172-178. MIT, 1978.

[9]
Gordon S. Novak Jr. and Donald Walker.
Multisource Tactical Information Integration.
Technical Report, SRI International, March, 1978.

[10]
Wolfgang Samlowski.
Case Grammar.
In E. Charniak and Y. Wilks, editor, _Computational Semantics_, . North-Holland, 1976.

[11]
Roger Schank and Robert Abelson.
Scripts, Plans, Goals, and Understanding.
Lawrence Erlbaum Associates, 1977.

[12]
Warren Teitelman.
INTERLISP Reference Manual.
Technical Report, XEROX Palo Alto Research Center, October, 1978.
section 17.16.

[13]
Ralph Weischedel.
Please Rephrase.
Technical Report, University of Delaware, 1976.

[14]
Yorick Wilks.
Good and Bad Arguments about Primitives.
Technical Report, University of Essex, 1976.

[15]
Terry Winograd.
Natural Language Understanding.
Academic Press, 1972.

[16]
William A. Woods.
Transition Network Grammars for Natural Language Analysis.
Communications of the ACM 13(10), 1970.