

WORD LEARNING WITH HIERARCHY-GUIDED INFERENCE*

David M. Keirse**

Department of Information and Computer Science
University of California, Irvine
Irvine, California

ABSTRACT

A technique for learning new words is discussed. The technique uses expectations generated by the context and an ISA hierarchy to guide the inference process. The learning process uses the context of several independent occurrences of the word to converge on its meaning.

I INTRODUCTION

Many of the words we know are not learned by having someone else tell us the definition or by looking them up in a dictionary; rather, we learn the words by extracting the meaning from the context. The learning of a word by context usually occurs over several instances. Several researchers [4, 8] have made attempts to learn new words from context, however, their programs did not try to extract the meaning over several occurrences of the word. Moreover, the first encounter with a new word usually gives only a vague meaning, because there are several possible interpretations. However, on successive encounters with the word, more can be learned about the word. The meaning becomes more refined by the addition of attributes extracted from context. Using the attributes to guide the selection of the definition of the word, we converge onto a known concept or idea that closely reflects the meaning. For example, consider the following three sentences, each having no connection between them other than the unknown word.

1. John picked up the foobaz.
2. John wrote with a foobaz.
3. John used an ink foobaz.

It should be clear from the three examples that the meaning of foobaz is closely related to a pen. Each sentence adds new information to the word foobaz, but the sentences alone are not sufficient to conclude the concept because foobaz could have been some other reasonable concept, such as: John wrote with a frenzy; John used an ink blotter.

Application of NLP techniques for building real prototype systems has a major stumbling block in that current NLP systems have handcrafted vocabularies. That is, each word must be known to the system a priori. This problem is especially evident in the Navy tactical message domain [6]. Navy communications send a large number of messages from ship to ship containing natural language text that report on their standard operations. Message understanding by machine is possible because the messages deal in a limited domain. However, the messages have many words that are acronyms, abbreviations, and generated names (e.g. call names), which an NLP system will not know at the time. Hence, a NLP system that learns words automatically is highly desired.

II REFINEMENT BY USING KNOWLEDGE

Expectation-based understanding [7, 2, 10] provides the system with information that can be useful in predicting the meaning of an unknown word [4]. These expectations are used as semantic restrictions on what should occur. When an unknown word is encountered, the expectations can serve as a hypothesis to the meaning. For instance, in one of its definitions, the verb phrase "picked up" has the semantic restriction on its direct object that it be a physical object. So in sentence one:

John picked up the foobaz.

a hypothesis could be that foobaz means "physical object." Physical object is a very general concept, and it is more likely the real intended meaning is more specific.

There are an infinite number of specializations of physical object that foobaz could be. The learning of the important attributes in refining

*This research was partially supported by grants N66001-81-C-0086 and N00123-81-1708 by the Naval Ocean Systems Center

**Now at Hughes Research Labs

the meaning of foobaz needs some guidance. General learning systems [9, 3] have had a problem of determining which set of features are important to generalize or differentiate on. The basic problem with the general learning approach is the lack of knowledge.

One form of knowledge representation useful in guiding the refinement of meaning is the ISA hierarchy. The ISA hierarchy, a well-known AI representation [1], provides an organization of concepts in which specialized concepts share properties of general concepts through inheritance, and also add properties of their own. The problem is to relate the unknown word with a known concept in the hierarchy. The hierarchy provides an organization for the merging of expectations gathered by context.

A. Hierarchy-guided deduction

Consider the following three sentences taken from three different Navy message reports of their activities.

1. Our aircraft overflew FRAMUS.
2. We saw running lights of FRAMUS.
3. Our ship overflown by aircraft of FRAMUS.

After sentence one, the inference can be made that FRAMUS is a SURFACE-CRAFT. The inference comes from the semantic restriction that the object can be overflown. That is, ships and subs are the only things that can be overflown in the Navy world. After sentence two, the FRAMUS should be a SHIP. This inference comes from the fact that ships are the only surface craft that have running lights, whereas submarines do not. The third sentence uses the fact that the carriers are only ships that possess aircraft. Notice that aircraft can be possessed by a carrier, a base, or a mission, so the restrictions gained by sentence one and two do restrict the possible meaning of FRAMUS in sentence three.

Figure 1 graphically represents the progressive refinement of the word FRAMUS: the numbers indicate the node FRAMUS was attached to after encountering each sentence. The ISA hierarchy is used to guide the refinement of the word meaning. Each descent of the tree adds more information to what the word means. That is, the lower nodes have more attributes, facts, and inferences attached to them, many of which are obtained by inheritance from higher concepts.

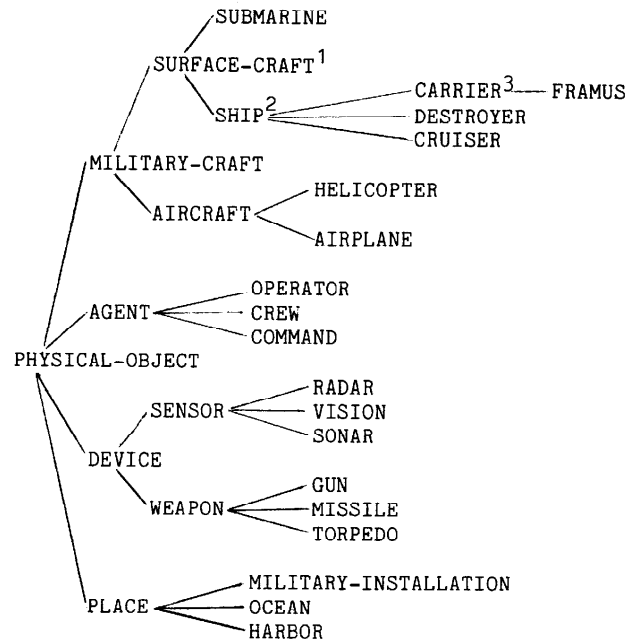


Figure 1: Hierarchy-guided Deduction

B. Tree Intersection

Suppose the encounters with the word happened in a different order. For example:

1. Our ship overflown by aircraft of FRAMUS.
2. Our aircraft overflew FRAMUS.
3. We saw running lights of FRAMUS.

The hierarchy provides a method of intersecting possible hypotheses. In this example the first hypothesis is that FRAMUS is either a carrier, a base, or a mission. The second message confirms that FRAMUS is a carrier, because the inference is that FRAMUS ISA ship from the semantic restriction of overflown and a carrier ISA ship. The other hypotheses are eliminated by the fact that neither a base or a mission ISA ship. The third message only confirms the hypothesis of FRAMUS being a carrier and does not add any more information. Changing the order in other ways eliminates hypotheses in different orders, but the result, FRAMUS ISA carrier, is the same.

C. Hierarchy-guided induction

Hierarchy guided induction, i.e., refining a word by traversing up the hierarchy tree, occurs when the initial guess of the meaning is more specific than the actual meaning. For example, take the following sentences:

1. TOI identified as Kashin.
2. TOI identified as a ship.
3. TOI identified as a surface craft.

(Kashin is a known word naming a particular class of Russian destroyer.)

Figure 2 illustrates the refinement of TOI by traversing up the tree.

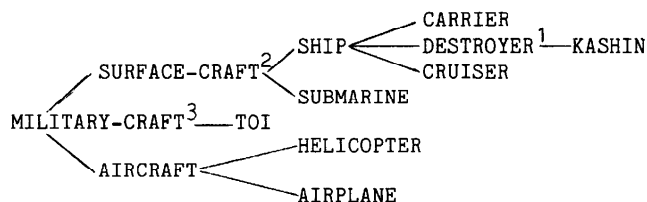


Figure 2: Hierarchy-guided Induction

From the first sentence it can be inferred that TOI is at least a one step above a Kashin (a Russian destroyer) in the classification hierarchy by noting that if TOI can be identified as a Kashin, then a TOI could be identified as another Russian destroyer other than a Kashin. With the same reasoning, TOI in the second sentence must be more general than a ship, for TOI could have been identified as a submarine, i.e., any surface craft. In the third sentence, likewise, a TOI must be some type of military craft, i.e., an aircraft or a surface craft. (TOI is an acronym for Target Of Interest.)

The use of both deduction and induction methods on a word could lead the hypothesized meaning to move up and down the tree indefinitely. The method of never moving a unknown word down the tree once the word has moved up the tree, prevents this vacillation from occurring. For example, if the sentence: "Diesel TOI detected" was encountered after the previous three sentences, then normally the rule can be used that only submarines are referred to as diesel, and therefore TOI must be a submarine. However, since TOI had been raised to MILITARY-CRAFT by induction, this rule cannot be invoked, because the rule would move TOI down the tree.

D. Detecting Errors

In the Navy tactical messages, the sentences are more often than not ungrammatical. Ungrammatical sentences can sometimes confuse the reader as to the usage of a word. Learning from an unknown word used ungrammatically can lead to eliminating all the hypotheses as to the meaning. Consider the following three sentences from three messages.

1. Carrier launched two HS helicopters.
2. HS gained contact on sub.
3. One helicopter from HS sighted hostile ship.

The hypotheses generated from sentence one are: HS is a type of helicopter, HS is a attribute of helicopter, or HS is some type of action that helicopters perform. Sentence two reduces the hypothesis to HS being a type of helicopter. Sentence three generates the expectations that HS is a ship, a base, a mission, or a task force. Since the intersection of current hypothesis and expectations is nil, there has been a mistake. As it turns out, the first sentence is grammatically incorrect. The correct sentence is:

Carrier launched two helicopters of HS.

HS actually stands for helicopter squadron, which is a task force of helicopters for the carrier. The expectations generated by the sentence are that HS is an adjective, noun modifying a noun, or a verb acting as a modifier, but not that HS is possessive, which is the proper syntactic construct needed in this instance. Thus, the incorrect expectations were generated. Unfortunately, there is no guarantee that syntactics errors will be detected; nevertheless, the technique has a chance to detect them.

E. Knowledge-based learning

This knowledge-based learning, using ISA hierarchies as the knowledge base, has the advantage that the knowledge can guide the learning process; the disadvantage is that most of the knowledge must be built in or acquired another way. The techniques being automatic (i.e., not interactive), they do not substantially build on the hierarchy tree. That is, the techniques only associate an unknown word to a concept already fully represented and do not try to form a totally new concept. However, other researchers have used hierarchies to do a form of knowledge acquisition [5], and a combination of the two methods could significantly add to the power of each.

III CONCLUSIONS

The ISA hierarchy is a useful representation that can guide the inference process in learning new words. This technique assumes that it takes several encounters with the new word to acquire a reasonable meaning of the word. It also assumes that the basic concept is already known.

This technique works well with most concrete nouns which can be easily represented in a hierarchy, but it does not work as well with word types that are not expressed as well in hierarchies, such as verbs. It is envisioned that other techniques of knowledge acquisition must be used in conjunction to acquire these words and concepts.

REFERENCES

- [1] Brachman, R. J.
On the Epistemological Status of Semantic Networks.
BBN Report 3807, Bolt Beranek and Newman Inc., April, 1978.
- [2] Cullingford, R.E.
Script Application: Computer Understanding of newspaper stories.
Technical Report, Yale University, 1978.
- [3] Dietterich, T.G. and Michalski, R.S.
Learning and Generalization of Characteristic Descriptions: Evaluation Criteria and Comparative review of selected methods.
In Proceedings of the Sixth International Joint Conference on Artificial Intelligence. International Joint Conference on Artificial Intelligence, August, 1979.
- [4] Granger, R.H., Jr.
FOUL-UP: A Program that figures out meanings of words from context.
In Proceedings of the Fifth International Joint Conference on Artificial Intelligence, pages 172-178. MIT, International Joint Conference on Artificial Intelligence, 1977.
- [5] Haas, N. and Hendrix, G. G.
An Approach to Acquiring and Applying Knowledge.
In First National Conference on Artificial Intelligence, pages 235-239. August, 1980.
- [6] Keirse, D. M.
Natural Language Processing applied to Navy Tactical Messages.
Technical Document 324, Naval Ocean Systems Center, February, 1980.
- [7] Schank, R.C.
Conceptual Information Processing.
North Holland, Amsterdam, 1975.
- [8] Selfridge, M.
A Computer Model of Child Learning.
In The First Annual Conference on Artificial Intelligence. AAAI, 1980.
- [9] Vere, S.A.
Induction of Relational Productions in Presence of Background Information.
In Proceedings of the Fifth International Joint Conference on Artificial Intelligence, pages 349-355. MIT, August, 1977.
- [10] Wilensky, R.
Understanding Goal-Based Stories.
Research Report 140, Department of Computer Science, Yale University, September, 1978.