Parsing and Representing Container Metaphors

R. Pascale
J. W. Roach
R. S. Virkar

Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

Abstract

We report the successful construction of a pattern based parser to recognize the class of container metaphors. Recognition of a metaphor in this class triggers a transformation that substitutes a correct, literal meaning form in the final representation of the utterance or sentence. The final meaning form reflects a theory of metaphors suggesting bodily experiences as the source of metaphor. A large set of primitives serves as the basic representation language. We conclude that pattern parsers with attached transformations work well for many normally difficult constructions such as metaphors, cliches and idioms.

I. Introduction

The frequency of metaphors in natural language has been reported to be as high as one in three utterances while some report occurrences in almost every utterance (Lakoff & Johnson, 1980). Recognizing, parsing and representing the meaning of a metaphor is therefore a major problem for any natural language parser.

Transformational syntactic parsers have little relevance to the problem, syntactic patterns of metaphors are not special, and normal methods of semantic analysis have not really found a method for determining the "real" meaning of a metaphor from its literal meaning. Syntactic patterns such as noun phrases, verb phrases and prepositional phrases do not allow a parser to distinguish between metaphorical and non-metaphorical uses of words. However, a parser that uses patterns of semantic primitives can detect whether or not a concept is used in its literal sense. The non-literal senses can thus be categorized as metaphorical.

In this paper, we present a parser based on patterns of primitives that can recognize and parse a class of metaphors known as container metaphors ("John is in love," for example). We also devise a means to represent the meaning of container metaphors appropriately.

II. What Are Metaphors?

By a metaphor we mean any non-literal use of a word or words. People normally use metaphors to express less concrete, less clear concepts, such as mental or emotional states, in terms of tangible concepts that are more easily visualized because of bodily experiences. The assimilation of concrete attributes by an abstract concept, however, must be only a partial structuring. A total assimilation of concrete properties would turn the abstract concept into a subcategory of a concrete object.

Metaphors tend to be cohesive. Orientational metaphors, for example, use direction and position in a (mostly) consistent manner to express meaning. "Down" metaphors have to do with lesser things or unhappy states while "up" metaphors refer to power and greater, happier concepts.

Metaphors such as orientational metaphors have their genesis in our culture and our experiences in the world. The fundamental theory we employ here, due to Lakoff and Johnson (1980), suggests that bodily experiences, for example, are responsible for the origin of metaphors. "Up is good" and "down is bad" then would derive from standing tall when we are happy and slouching when we are depressed. Being "in love" is having love surround and engulf our thoughts.

Lakoff and Johnson (1980) classify metaphors into four categories: orientational, structural, ontological and imaginative. Ontological metaphors express events, activities, emotions and other abstract concepts as entities and substances. An example of this category of metaphors is "mind is a machine." This metaphor gives rise to sentences such as "I am a little rusty today," and "My mind just isn't operating today." This type of metaphor implies that an abstract object and some other object to which it is being compared have the same qualities.

The ontological metaphors serve the purposes of referring (e.g. "That was a beautiful catch"), quantifying (e.g. "Dupont has a lot of power in Delaware"), identifying aspects (e.g. "The brutality of war dehumanizes us all"), identifying causes (e.g. "He did it out of anger"), and setting goals and motivating actions (e.g. "He went to New York to seek fame and fortune"). There are many sub-classes of ontological metaphors. These sub-classes are container metaphors, personifications and metonymy.

In the sub-class of container metaphors, each concept has an in-out orientation, bounding surface, container object, substance and other qualities. Some of these are land areas (e.g. "There is a lot of land in Kansas"), visual field (e.g. "He is in my view") and states (e.g. "He is in love"). In the sub-class of personifications inanimate objects are allowed to possess human qualities. An example of such a metaphor is "The feather was dancing in the wind." The feather (an inanimate entity) is given the human quality of dancing. The sub-class of metonymy is similar to personifications, however, in metonymy one entity is substituted for another. An example of such a metaphor is "The sax is out sick today." Here, the sax actually refers to the person who plays the saxophone.

In this paper, we shall concentrate on the container metaphors. Container metaphors are often used by people without fully being aware of the non-literal sense of the words. Examples of this sub-class are abundant in written text as well as spoken language.

III. Pattern-based Parsing

The use of a pattern-based approach for parsing natural language input was first exploited on a large scale by Parkison, Colby and Faught (1977). Their approach entailed matching pieces of input to elements of a large base of prestored patterns, and successful matches resulting in changes in the original sentence such as simplification and replacements. It has been shown since that purely syntactic means are not sufficient to relate meaning to utterances (Gross, 1979). We have already used pattern-based parsing approach in conjunction with semantic primitives (Virkar & Roach, 1988a, 1988b; Sanford & Roach, 1988). Now, we shall describe our approach for metaphors.

Linguistic expressions are generated from the lexicon and we hypothesize that most of the every-day lexicon can be represented by a large, yet finite, number of semantic primitives. Sentence forms can be classified by the patterning of semantic primitives and function words. We hypothesize that all simple sentence forms can be captured by a very large, yet finite, set of patterns.

A semantic primitive represents a set of words that refer to the same concept. Concepts interact with each other and these relationships are seen through the expressions of natural language. Our classification has four basic classes of semantic primitives: events, entities, abstracts and relationals. A meaningful expression is an expression that describes an event with the help of other associated basic classes. To describe an event, one or more of these associated elements may not be required. A meaningful expression cannot be anomalous, indeterminable, nor contradictory (Allan, 1986).

If we denote the set of all semantic primitives by P. Then, P^+ denotes the set of all sequences, of length ≥ 1, of semantic primitives. The language L, the set of all meaningful expressions, can now be viewed as a subset of the set of all possible sequences derived from P. In other words, $L \subset P^+$. Natural language understanding can now be defined as a mapping, U, that translates all meaningful expressions in L onto the set of sentence meaning structures, S; i.e., $U: L \rightarrow S$.

We use semantic mappings to translate an utterance, based on its semantic pattern, into a possible interpretation. Mappings use axioms to eliminate the incorrect interpretations of an utterance. We say that a function word is a word that, in its position in the utterance, signals the beginning of a new (possibly primitive) meaningful expression. The set of function words F contains logical connectives such as 'and' and 'or', and prepositions such as 'for', 'of', etc. It is also important to note that two meaningful expressions can be connected without the presence of a function word when one expression is embedded in another in a subordinate fashion.

A primitive utterance is a sentence that conveys only one "meaningful expression" in a language. We shall call it a primitive meaningful expression. A complex utterance, on the other hand, is a sentence that expresses two or more (related) meaningful expressions connected by one or more function words. A simple complex utterance is one that has two primitive meaningful expressions connected by one function word. Hence, by recognizing the function word, the two expressions can be transformed into two primitive meaningful expressions.

A semantic transformation can be defined as the process of decomposing a complex utterance into two primitive meaningful expressions using the function word appearing in the complex utterance, without altering the 'meaning' expressed by the original complex utterance. It should be noted, however, that several semantic transformations may exist for a given function word, and the one that is applied is selected based on the semantic pattern of the sentence. It should also be noted that a semantic transformation on an utterance that does not contain any function words is equivalent to applying a semantic mapping.

It is possible that after applying a semantic transformation, one or both of the resulting sentences will be complex. This can occur only in the presence of more function words. These function words can now be used to apply other semantic transformations and further reduce the sentences. Since every execution of a semantic transformation reduces the complexity of the sentence, a finite number of semantic transformations guarantees convergence to primitive sentences.

IV. Representation

In the previous section we have described the basic steps in the working of the parser. We acknowledged that multiple interpretations exist for words within the context of an utterance. The problem of resolving these ambiguities becomes evident in the treatment of metaphors.

The representation language we use is based on a set of semantic primitives. This set has been adapted from a linguistic effort (Nida, 1975) of developing a thesaurus. The parsing of an utterance produces a meaning form consisting of a many-sorted representation. This representation involves the event and other components, namely entities, relationals and abstracts.

In the case of container metaphors, the containing object is not physically (or spatially) containing some other object. Thus, a relational form such as

SPATIAL :: < object1 > < object2 >

is improper and hence must be discarded. If the containing concept is an emotion, then the parser should represent the affecting state of experience. If the containing concept is an activity or a state of affairs, then the parser should show the involvement. Similarly, if the containing concept is a class of objects, then the parser should represent the membership. Based on the type of the container metaphor our parser produces a representation that conveys appropriate meaning.

A container metaphor involving 'time' allows expressions such as 'in an hour' and 'in seconds'; and our representation scheme produces

{ Event X } :- Abstract (DURATION 1 hour)
{ Event X } :- Abstract quickly

respectively. Thus, the container-like use of time is deciphered correctly where time is actually an abstract primitive that describes the duration of events.

A container metaphor involving social groups allows expressions such as 'in a fraternity' and 'in computer science'; parsing of such metaphorical expressions results in

Membership :: < Entity Y > < Social Group fraternity >
Membership :: < Entity Y > < Profession comp sci >

respectively. The groups expressed as containers can be represented by using the membership relational.

A container metaphor involving emotional states provides expressions such as 'in love' and 'in pain.' The representation scheme we employ produces the following form for such expressions

Experience State love [< Entity Y >]

Thus, the container-like use of experiential states can be represented by states of entities.

A container metaphor with activities and events gives rise to expressions such as 'in the race' and 'in Watergate.' Our representation scheme generates the following form for these expressions

Involvement :: < Entity X > < Event race >

Involvement is a relational that captures emotional, physical as well as conceptual entanglement. Our scheme represents it appropriately and the event, used as a container, is shown to involve the entity X. It should be noted that this Involvement relational has nothing in common with the constraint relation defined by Barwise and Perry in situation semantics (Barwise & Perry, 1983).

Our representation scheme is based on four classes of semantic primitives. This allows it to differentiate between different types of container metaphors. As can be seen from the examples above, the trigger word for the container metaphors, "in", can relate an entity to an affecting relational, an affecting event, or an affecting abstract.

V. Results

We found that there was a one-to-one correspondence between sub-classes of container metaphors and transformation rules. Every container metaphor that we could figure out was correctly parsed by the rules we constructed. We identified fifteen sub-classes of container metaphors and added transformation rules for these sub-classes. These transformation rules can parse a large number of sentences with container metaphors. The rules we constructed are based on patterns of semantic primitives, and as such, each rule accounts for all utterances that fit the pattern associated with that rule.

Obviously, we cannot guarantee completeness of the rules, but for any example not covered by the rule, there will be no difficulty adding a pattern and its associated transformation to the rule base. Table 1 contains a sampling of the sentences that our system can handle.

VI. Discussion

The system we have built is one of the few computational linguistics systems to take prepositions seriously. By that we mean that our system can parse a very large number of word senses (captured as patterns) for each preposition. The only other previous system to our knowledge to work seriously with prepositions studied only the word 'for' (Hemphill, 1981). No previous system, for example, has attempted to capture over one hundred senses for the word 'over' (data source: Brugman, 1981) or any of the other prepositions. In fact, few systems have really attempted to deal with the polysemy problem at all. Most of our data for the voluminous number of prepositional word senses comes from Hill (1968). Our system derives its power to recognize metaphors precisely from this ability to account for the numerous senses of prepositions. To the extent that metaphorical use of language can be associated with prepositional phrases, our system can handle metaphors.

Container metaphors account for only one set of preposition triggered metaphorical phrases; we expect to extend our work to other prepositional metaphorical phrases. We hypothesize that pattern based parsers based on a large set of primitives and designed to help solve the polysemy problem will help solve the metaphor recognition problem. Solving the metaphor representation problem, of course, requires a theoretical stance, such as the one put forth in Lakoff and Johnson (1980).

VII. Conclusions

This paper has presented a parsing technique based on patterns of primitives that can recognize and parse metaphorical container phrases. The representation techniques used here reflect a theory of metaphors that requires metaphorical expressions to originate in bodily experiences. Experiments with the system indicate excellent results for the class of metaphors that were the target. Extensibility of the parsing techniques depends on the applicability of pattern based parsing to the recognition of metaphorical structures. We hypothesize that a large set of metaphorical structures can be parsed using pattern based techniques.

References

Allan, K. 1986. Linguistic Meaning: Vol. 1. New York: Routledge and Kegan Paul.

Barwise, J. and Perry, J. 1983. Situations and Attitudes. Cambridge, Massachusetts: The MIT Press.

Brugman, C. 1981. "Story of Over," Master's Thesis, University of California, Berkeley.

Gross, M. 1979. "On the Failure of Generative Grammar," Language, vol. 55, no. 4, pp. 859-885.

Hemphill, L. G. 1981. "A Conceptual Approach to Automated Language Understanding and Belief Structures: With Disambiguation of the word 'For',"Dissertation submitted to the Department of Computer Science, Stanford University.

Hill, L. A. 1968. Prepositions and Adverbial Particles. London: Oxford University Press.

Lakoff, G. and M. Johnson. 1980. Metaphors We Live By. Chicago: The University of Chicago Press.

Nida, E. A. 1975. Componential Analysis of Meaning: An Introduction to Semantic Structures. The Hague: Moulton.

Parkison, R. C., K. M. Colby and W. S. Faught. 1977. "Conversational Language Comprehension Using Integrated Pattern-Matching and Parsing," Artificial Intelligence Journal, vol. 9, no. 2, pp. 111-134.

Sanford, D. L. and J. W. Roach. 1988. "A Theory of Dialogue Structures to Help Manage Human-Computer Interaction," IEEE Transactions on Systems, Man and Cybernetics- Special Issue on Human-Computer Interaction and Cognitive Engineering, vol. 18, no. 4 (July/August), pp. 567-574.

Virkar, R. S. and J. W. Roach. 1988a. "Pattern-Based Parsing for Word Sense Disambiguation," Proceedings of The Tenth Annual Conference of the Cognitive Science Society, pp 688-694.

Virkar, R. S. and J. W. Roach. 1988b. "Direct Assimilation of Expert-Level Knowledge by Automatically Understanding Research Paper Abstracts," International Journal of Expert Systems, vol. 1, no. 4, pp. 281-305.

Table1. Expressions classified by what is being metaphorically represented

1. TIME

a. She ran the mile in five minutes.

RUSH-MOTION [{<PERSON>}:- DISTANCE
{<PERSON>}:- DURATION]

2. EMOTION

a. Fred is in love.

EXPERIENCE STATE love [<PERSON>]

b. She is in a panic.

EXPERIENCE STATE panic [<PERSON>]

3. GEOGRAPHICAL AREAS

a. The dog is in the field.

POSITION:: [<ANIMAL> <PLACE field>]

b. The tree is in the yard.

POSITION:: [<PLANT> <PLACE yard>]

c. The house is in Delaware.

POSITION:: [<DWELLING> <PLACE Delaware>]

4. SOCIAL GROUPS

a. He is in a fraternity.

MEMBERSHIP:: [<PERSON> <SOCIAL GROUP fraternity>]

b. She is in biology.

MEMBERSHIP:: [<PERSON> <PROFESSION biology>]

5. EVENTS and ACTIVITES

a. He was in Watergate.

INVOLVEMENT:: [<PERSON> <EVENT Watergate>]

b. She is in the race.

INVOLVEMENT:: [<PERSON> <EVENT race>]