Semantic Interpretation Using KL-ONE

Abstract

This paper presents extensions to the work of Bobrow and Webber [Bobrow&Webber 80a, Bobrow&Webber 80b] on semantic interpretation using KL-ONE to represent knowledge. The approach is based on an extended case frame formalism applicable to all types of phrases, not just clauses. The frames are used to recognize semantically acceptable phrases, identify their structure, and relate them to their meaning representation through translation rules. Approaches are presented for generating KL-ONE structures as the meaning of a sentence, for capturing semantic generalizations through abstract case frames, and for handling pronouns and relative clauses.

1. Introduction

Semantic interpretation is the process of relating the syntactic analysis of an utterance to its meaning representation. Syntactic analyses associate immediate constituents with their syntactic function in a matrix constituent, e.g., the sentence "Send him the message that arrived yesterday.", has a syntactic analysis in RUS [Bobrow 78] as shown in Figure 1. The elements of the meaning representation are the objects, events, and states of affairs perceived by the speaker. The relationships between these entities will be called semantic functions. The basis for our semantic processing scheme is a familiar one based on that of case frames used to describe clause structure [Bruce 75]. Our case frames are used for all phrase types: clauses, noun phrases, prepositional phrases, etc. We choose to represent both the syntactic and semantic analyses in the knowledge representation language KL-ONE [Brachman&Schmolze 82, Schmolze&Lipkis 83, Moser 83]. The essential properties for the meaning representations constructed are that each concept represents a semantic constituent and each of its roles identifies the semantic function of one of its immediate constituents. Figure 2 gives an analysis of the example sentence above. We have picked a constituent structure and names for semantic functions fitting the computer mail application of the Consul project at USC/Information Sciences Institute [Kaczmarek 83]. The exact details of the analysis are not critical; the essential point is that

1This material is based upon work supported in part by the Defense Advanced Research Projects Agency under Contract Numbers MDA 903-81-C-0335, ARPA Order No. 2223, and N00014-77-C-0378, ARPA Order No. 3414. Views and conclusions contained in this paper are the authors' and should not be interpreted as representing the official policies of DARPA, the U.S. Government, or any person or agency connected with them.

2We use this sentence to illustrate many of the points in this paper. Assume that "yesterday" modifies "arrived."

3All of the KL-ONE diagrams in this paper are simplified for expository purposes.
Our framework does not assume that a syntactic analysis of a complete sentence is found before semantic interpretation begins. Rather, the implemented semantic interpreter proceeds incrementally as the grammar proposes the syntactic function of an immediate constituent; this model of communication between syntax and semantics has been termed a cascade [Woods 80, Bobrow & Webber 80b].

To achieve semantic interpretation, some well-known types of knowledge need to be employed, e.g., selection restriction information (often represented using semantic features), structural information (often encoded in case frames), and translation information (often defined with various kinds of projection rules).

Some of the difficulties in representing and applying this knowledge include the following:

1. Translation rules (projection rules) for generating correct meaning representations must be defined. We have been able to define modular projection rules that make use of the inheritance properties of KL-ONE.

2. Since much of the knowledge for a particular application is necessarily domain specific, it is important to organize it in a way to ease extension of the knowledge base and to ease moving to a new domain.

3. Since distributional restrictions require specific semantic features, pronouns and other semantically neutral terms not necessarily having those features must be accepted wherever they are consistent with the expected type of noun phrase.

4. The inter-constituent relationships arising in relative clauses must be consistent with all selection restrictions and be represented in the resulting meaning representation.

This paper addresses each of these issues in turn.

We are building on techniques presented by Bobrow and Webber [Bobrow & Webber 80a, Bobrow & Webber 80b]. This paper describes the system currently in use at USC Information Sciences Institute. The basic framework is reviewed in Section 2. Section 3 presents the translation mechanism [Sondheimer 84]. Capturing semantic generalizations is the topic of Section 4. Sections 5 and 6 discuss issues regarding pronouns and relative clauses, respectively. Related work is identified in Section 7. The final section summarizes the results, and identifies further work. A very brief introduction to KL-ONE is provided in an appendix.

2. Background

The framework being developed uses a frame for each semantically distinguishable type of phrase. Thus, a frame will be required for each class of phrase having a unique combination of

- semantic distribution,
- selection restrictions on constituents making up the phrase, and
- assignment of semantic relations to syntactic function.

It is likely that the frames will reflect the natural categories of descriptions of objects, events, actions, and states of affairs in any particular application. For example, in the computer mail domain, the following are some frames that have been useful:

- Clauses describing the sending of messages: SEND-CLAUSE
- Clauses describing message arrival: ARRIVE-CLAUSE
- Noun phrases describing messages: MESSAGE-NP
- Noun phrases describing senders and recipients: USER-NP

In the framework developed by Bobrow and Webber [Bobrow & Webber 80a, Bobrow & Webber 80b], for each frame, each possible immediate constituent is associated by syntactic function with a case or slot. The clause frames have slots identified as head, subject, direct object, indirect object, etc. Noun phrase frames have slots for the head, adjective modifiers, article, etc. Each slot specifies the fillers that are semantically acceptable, whether it is required or optional, and the number of times it may be filled in a phrase. The constraints on fillers of frames' slots are stated in terms of other frames, e.g., the direct object of a SEND-CLAUSE must be a MESSAGE-NP, or in terms of word senses and categories of these senses. Some example word sense categories are:

- Message description nouns, such as "message" or "letter": MESSAGE-NOUN
- Information transmission verbs, such as "send" or "forward": TRANSMISSION-VERB

In our domain the constraint on the subject of an ARRIVE-CLAUSE is that it satisfies the MESSAGE-NP frame. A constraint on the head of the MESSAGE-NP frame is that it is a word sense in the category MESSAGE-NOUN.

Frames are represented as KL-ONE concepts. Case slots appear as roles of concepts. Semantic constraints on what can fill a case slot are encoded as the value restrictions of roles. These value restrictions are concepts representing frames, word senses, or word sense categories. Number restrictions on roles show the number of times the syntactic function may be realized. A required slot is marked by the number restriction on its role having a minimum of 1; an optional slot has a number restriction with a minimum of 0 and a maximum greater than 0. A phrase is said to instantiate a given frame X if and only if its immediate constituents satisfy the appropriate value and number restrictions of all of X's roles. The collection of frames and word sense

\[\text{assignment of semantic relations to syntactic function.}\]

It is likely that the frames will reflect the natural categories of descriptions of objects, events, actions, and states of affairs in any particular application. For example, in the computer mail domain, the following are some frames that have been useful:

- Clauses describing the sending of messages: SEND-CLAUSE
- Clauses describing message arrival: ARRIVE-CLAUSE
- Noun phrases describing messages: MESSAGE-NP
- Noun phrases describing senders and recipients: USER-NP

In the framework developed by Bobrow and Webber [Bobrow & Webber 80a, Bobrow & Webber 80b], for each frame, each possible immediate constituent is associated by syntactic function with a case or slot. The clause frames have slots identified as head, subject, direct object, indirect object, etc. Noun phrase frames have slots for the head, adjective modifiers, article, etc. Each slot specifies the fillers that are semantically acceptable, whether it is required or optional, and the number of times it may be filled in a phrase. The constraints on fillers of frames' slots are stated in terms of other frames, e.g., the direct object of a SEND-CLAUSE must be a MESSAGE-NP, or in terms of word senses and categories of these senses. Some example word sense categories are:

- Message description nouns, such as "message" or "letter": MESSAGE-NOUN
- Information transmission verbs, such as "send" or "forward": TRANSMISSION-VERB

In our domain the constraint on the subject of an ARRIVE-CLAUSE is that it satisfies the MESSAGE-NP frame. A constraint on the head of the MESSAGE-NP frame is that it is a word sense in the category MESSAGE-NOUN.

Frames are represented as KL-ONE concepts. Case slots appear as roles of concepts. Semantic constraints on what can fill a case slot are encoded as the value restrictions of roles. These value restrictions are concepts representing frames, word senses, or word sense categories. Number restrictions on roles show the number of times the syntactic function may be realized. A required slot is marked by the number restriction on its role having a minimum of 1; an optional slot has a number restriction with a minimum of 0 and a maximum greater than 0. A phrase is said to instantiate a given frame X if and only if its immediate constituents satisfy the appropriate value and number restrictions of all of X's roles. The collection of frames and word sense

\[\text{assignment of semantic relations to syntactic function.}\]
information is called a *Syntaxonomy* (for syntactic taxonomy), since it encodes knowledge regarding semantic interpretation in a hierarchy of syntactic classes.

3. Translation Rules

To achieve the mapping from syntactic analysis to meaning representation, translation rules are associated with individual frames. Though the rules we give generate KL-ONE structures as the meaning representation, other translation rules could be developed for generating forms in a different target representation language.

Any KL-ONE concept C representing a frame has an associated concept C' representing the main predicate of the translation. For example, the translation of SEND-CLAUSE is the concept Send-mail. Translations are stored in data attached to the frame; we label this data TRANSLATION.

The translation rules themselves can be associated with individual case slots. When inheritance results in more than one translation rule for a case slot, the one originating from the most specific frame in the hierarchy is selected.7

Suppose we are building the translation C' of a matched frame C. One common translation rule that could appear at a role R of C is (Paraphrase-as R'). This establishes the translation of the filler of R as the filler of R' at concept C'. For example, the indirect object slot of SEND-CLAUSE has the rule "(Paraphrase-as addressee)" to map the translation of the noun phrase in the indirect object position into the addressee role of the Send-mail.

Another rule form, (Attach-SD sf), takes a semantic function sf as an argument and attaches the translation of the constituent filling R as the filler F of sf. A example of its use in the processing of relative clauses as described in Section 6. Attach-SD differs from Paraphrase-as by having facilities to establish a role from F to C'. This automatic feature is essentially the opposite of Paraphrase-as, in that a semantic function runs from the embedded constituent to its matrix phrase.

Another rule form is not a translation rule per se, but stores data with the syntactic concept representing the syntactic analysis of the phrase. The data could be checked by other (conditional) translation rules.

Underlying these forms and available for more complex types of translation is a general mechanism having the form "source = > goal." The source identifies the structure that is to be placed at the location identified by the goal. The formalism for the source allows reference to arbitrary constants and concepts and to a path through the concepts, roles, and attached data of a KL-ONE network. The goal formalism also shows a path through a network and may specify establishment of additional roles.

A separate test may be associated with a translation rule to state conditions on the applicability of a rule. If the test is false, the rule does not apply, and no translation corresponding to that role is generated. The most common type of condition is (Realized-Function? role), which is true if and only if some immediate constituent fills that role in the analysis. It can be used as an explicit statement that an optional role is translated only if filled or as a way of stating that one constituent's translation depends on the presence of another role. Additional conditions are (EMPTY-ROLE? role), which checks that role is not filled, and (ROLE-FILLER? role class), which checks that the filler of role is of type class. Since all three take a role name as argument, they may be used to state cross-dependencies among roles.

Figure 3 contains some of the frames that allow for the analysis of our example. The treatment of the pronoun and relative clause in the example sentence of Section 1 will be explained in Sections 5 and 6.

4. Capturing Semantic Generalizations via Abstract Case Frames

Verbs can be grouped with respect to the cases they accept [Simmons 73, Celce-Murcia 76, Gawron 83]; likewise, groups exist for nouns. A KL-ONE syntaxonomy allows straightforward statement of common properties, as well as individually distinct properties of group members. Abstract case frames are semantic generalizations applicable across a set of the familiar sort of concrete frames. Properties common to the generalization can be defined at the abstract frames and related to the concrete frames through inheritance.

The use of time modification in "that arrived yesterday" is the same as that of other verbs describing completion of an activity, e.g., "come", "reach", and "finish". A general frame for clauses with these verbs can show this role. The concrete frames for clauses with verbs in this group are subconcepts and thereby accept the time modifier (see Figure 4). The concrete frames can restrict both the number and type of time modifiers, if necessary. Translation rules associated with this time role can also be restricted at the concrete frames.

Some modifiers dramatically affect the translation of entire phrases, as in the partitive modifier "half of". A description of "half of some individual entity (as opposed to a set of entities) may not have the same distribution. For example, "Delete this message from my directory.", makes sense, but "Delete half of this message from my directory.", does not. This can be easily stated through an abstract frame for the basic message description specialized by two concrete frames (see Figure 5).

A related case is "toy X." The translation of "toy X" is certainly different from that of X, and their distributions may differ as well. This may be handled in a way similar to the partitive example.8 This class of examples points out the limits of case frame systems. Other modifiers, such as "model" and "fake", are easily recognizable. However, more complex modifiers also make the same distinctions, e.g., "The gun that was a fake was...

---

7 There is also an escape mechanism that allows inheritance of all rules not indexed to any role.

8 An interesting alternative is to show the toy modifier as an optional role on an abstract frame for object descriptions. Underneath it could be an abstract frame distinguished only by requiring the toy modification role. All appropriate inferences associated with descriptions of toys could be associated with this concept. Frames for the basic descriptions of specific object types could be placed underneath the object description frame. These could recognize "toy X." Our systems invoke the KL-ONE classifier after the recognition of each phrase [Schmolze & Lipkis 83]. In this case, classification will result in identification of the phrase as a kind of both X description and toy description allowing translation to show what is known about both without creating a "toy X" frame by hand. We have not completely analyzed the affect of this strategy on the translation system.
TRANSLATION: "Send-mail"

ARRIVE-CLAUSE

TRANSLATION: "Arrival-mail"

MESSAGE-NP

TRANSLATION: "ComputerMail"

Figure 3: Some frames used for "Send him the message that arrived yesterday."

Figure 4: A fragment of the syntaxonomy. Double arrows are subc relationships, i.e., essentially "is-a" arcs. Not all roles are shown.

Figure 5: Syntaxonomy for partitives.

104
John’s.” and “The gun that was made of soap was John’s.”. Viewing our semantic interpretation system as a special purpose inference system, it seems prudent to leave the recognition of the type of these “guns” to more general-purpose reasoners.

Abstract case frames have significantly eased the development and expansion of semantic coverage within our application by helping us to focus on issues of generality and specificity. The new frames we add have many slots established by inheritance; consistency has been easier to maintain; and the structure of the resulting syntaxonomy has helped in debugging.

5. Semantically Neutral Terms

Case frames are an attempt to characterize semantically coherent phrases, for instance, by selection restrictions. In computational linguistics, selection restrictions have been applied to the constituents that are possible fillers rather than what the constituents denote. For example, the restriction on the direct object of a SEND-CLAUSE is MESSAGE-NP, rather than messages. Problems with using such approximations in parsing are discussed in [Ritchie 83].

For many natural language interfaces, a noun phrase’s internal structure gives enough information to determine whether it satisfies a restriction. However, there are forms whose semantic interpretation does not provide enough information to guarantee the satisfaction of a constraint and yet need to be allowed as fillers for slots. These include pronouns, some elliptical forms, such as “the last three”, and other neutral noun phrase forms, such as “the thing” and “the gift”. This also includes some nonlexical gestural forms like the input from a display that shows where the user pointed (literally or via a mouse). We refer to all of these as semantically neutral terms. A semantic interpretation system should accept such forms without giving up restrictions on acceptable semantic categories. However, these forms cannot, in general, appear everywhere. In discussing computer mail, “I sent him” should be considered nonsense.

Bobrow and Webber [Bobrow&Webber 80b] propose a general strategy for testing the compatibility of a constituent as a slot filler based on non-incompatibility. The current system at USC/ISI takes a conservative view of this proposal, developing the idea for only neutral reference forms. All noun phrase types displaying neutral reference are defined as instances of the concept NeutralReference-NP. Furthermore, disjointness relations are marked between the various subclasses of neutral references and those classes of explicit descriptions which have nonintersecting sets of potential references. During interpretation, when such a NeutralReference-NP is proposed as a slot filler, and that concept is not disjoint from the value restriction on the slot, it is accepted.

In addition, since the slot restriction and the filler each have meaning of their own, e.g., “he” describes a human male in the computer mail domain, the translation should show the contribution of both the neutral term and the constraint on the slot. When the neutral form is qualified as a constituent by the system, both the neutral form and the selection constraint are remembered. When it is time to produce the translation, the translation rule for the slot applies to a concept which is the conjunction of the translations of the neutral reference form and the restriction.

Part of the network that supports the translation of “he” in the example of section 1 is shown in Figure 6. Referring to Figures 2 and 3, the effect of a reference to a male where a reference to a computer-user was expected can be seen.

6. Inter- Constituent Relationships: Relative Clauses

In relative clauses, the constraint on the slot filled by the relative pronoun or the trace must be satisfied by the noun phrase that the relative clause modifies. In addition, the translation of the noun phrase must reflect the contribution of the use of the pronoun or trace in the relative clause. For example, in “Send him the message that arrived yesterday”, the constraint on the subject of “arrive” must be satisfied by the noun phrase of which it is a part. Further, translation must result in co-reference within the meaning representation of the value of the message role of the ARRIVAL-MAIL concept and the value of the message role of the Send-mail concept (see Figure 2). This is a form of inter-constituent relationship.

Our system processes relative clauses by treating the relative pronouns and trace elements as neutral reference forms (just as in the pronominal cases discussed in Section 5 and by storing the constraints on the head of the relative clause until they can be employed directly. In our example, the noun phrase “that” is seen as a Trace-NP, a kind of NeutralReference-NP. The structure assigned “that” is compatible with MESSAGE-NP and hence acceptable. On translation, the Trace-NP is treated like a neutral reference but the role and unchecked constraint are recorded as attached data on the instantiated case frame that results from parsing the arrival clause. In the example, the facts that a Trace-NP is in the subject role and that a Message-NP is required are stored. That constraint is tested against the classification of the matrix noun phrase when the clause is proposed as a relative clause modifier.11

Figure 6: Network for “he.” Note that computer User is a subconcept of Person.

10 The RUS parser which we employ supplies a “trace” to establish a syntactic place holder with reduced relatives.

11 If the use of the relative pronoun or trace is inside a phrase inside the relative clause, as in “the town from which I come”, the role and constraint will be passed upward twice.
If that constraint is satisfied, the fact that the relative pronoun and noun phrase co-refer is recorded. When the entire noun phrase is processed successfully, the appropriate co-references are established by performing (Attach-SD further-constraint) and by retrieving the translation associated with the role filled by the Trace-NP. This establishes co-reference between the concept attached by the translation rule and the translation of the entire noun phrase. In our example, the translation of the noun phrase is made the value of the message role of the Arrival-mail.

7. Related Work

Our technique uses properties of KL-ONE to build a simplified, special-purpose inference engine for semantic interpretation. The semantic processor is separate from both syntactic and pragmatic processing, though it is designed to maintain well-defined interaction with those components through Woods' cascade model of natural language processing [Woods 80]. Uniform methods include logic grammars [Pereira 83, Palmer 83] and semantic grammars [Burton 77, Hendrix 78, Wilensky 80]. Logic grammars employ a Horn-clause theorem prover for both syntactic and semantic processing. Semantic grammars collapse syntactic and semantic analysis into an essentially domain-specific grammar. Semantic interpretation is handled through unification in some evolving systems, such as PATTR-II [Robinson 83].

Several recent systems have separate semantic interpretation components. Hirst [Hirst 83] uses a Montague-inspired approach to produce statements in a frame language. He uses individual mapping rules tied to the meaning-affecting rules of a grammar. Boguraev [Boguraev 79] presents a semantic interpreter based on patterns very similar to those of our case frames. The meaning representation it produces is very similar to the structure of our case frames.

8. Conclusion

We have presented approaches to typical difficulties in building semantic interpreters. These have included a sketch of a translation system that maps from the matched frames to KL-ONE meaning representations. The idea of abstract case frames and applications of them were introduced. Finally, ways of accepting neutral references and allowing for the inter-constituent constraints imposed by relative clauses were presented.

Our experience indicates that KL-ONE is effective as a means of building and employing a library of case frames. The basic approach is being used in research computer systems at both USC/Information Sciences Institute and Bolt Beranek and Newman, Inc.

Of course, many problems remain to be solved. Problems currently under investigation include:

- Feedback from the pragmatic component to guide semantic interpretation,
- Generation of error messages (in English) based on the case frames if the request seems beyond the system's capabilities,
- Understanding classes of metonymy, such as "Send this window to Jones," and
- Provision for meaningful use of nonsense phrases, such as "Can I send a package over the ARPAnet?"

I. Brief Description of KL-ONE

KL-ONE offers a rigorous means of specifying terms (concepts) and basic relationships among them, such as subset/superset, disjointness, exhaustive cover, and relational structure. Concepts are denoted graphically as ovals. Concepts are structured objects whose structure is indicated by named relations (roles) between concepts. Roles are drawn as arcs containing a circle and square. The concepts at the end of the role arcs are said to be value restrictions. In addition, roles have maximum and minimum restrictions on the number of concepts that can be related by the role to the concept at the origin of the arc. Concepts can also have data attached to them, stored as a property list. Finally, the set of concepts is organized into an inheritance hierarchy, through sub relations drawn with double-line arrows from the subconcept to the superconcept.

All of the KL-ONE diagrams in the text are incomplete; for instance, Figures 3 and 5 focus on different aspects of what is one KL-ONE structure. In figure 3, the diagram for Send-CLAUSE specifies the concepts of "send" clauses. They have exactly one head, which must be the lexical concept "send." They must have a direct object which is a MESSAGE-NP, and they optionally have an indirect object which is a USER-NP. Figure 5 shows that Send-CLAUSE's are MESSAGE-TRANSMISSION-CLAUSE's, which are a type of CLAUSE.

The meaning representation, Figure 2, generated for "Send him the message that arrived yesterday" consists of the concept Send-mail, having an addressee which is a Computer-User and a message which is ComputerMail.

References

[Bobrow 78] R.J. Bobrow, "The RUS System," in B.L. Webber, R. Bobrow (eds.), Research in Natural Language Understanding, Bolt, Beranek, and Newman, Inc., Cambridge, MA, 1978. BBN Technical Report 3878.

[Bobrow&Webber 80a] Robert Bobrow and Bonnie Webber, "PSI-KLONE: Parsing and Semantic Interpretation in the BBN Natural Language Understanding System," in Proceedings of the 1980 Conference of the Canadian Society for Computational Studies of Intelligence, CCSI/SCEIO, May 1980.
