ON KNOWLEDGE REPRESENTATION USING
SEMANTIC NETWORKS AND SANSKRIT

S.N. Srihari, W.J. Rapaport, D. Kumar

87-03 February 1987

Department of Computer Science
State University of New York at Buffalo
226 Bell Hall
Buffalo, New York 14260

This work was supported in part by the National Science Foundation under grant no. IST-8504713, and in part by the Air Force Systems Command, Rome Air Development Center, Griffiss Air Force Base, NY 13441 and the Air Force Office of Scientific Research, Bolling AFB, DC 20332 under contract no. F30602-85-C-0008.
ON KNOWLEDGE REPRESENTATION USING SEMANTIC NETWORKS
AND SANSKRIT

Sargur N. Srihari, William J. Rapaport, and Deepak Kumar

Department of Computer Science
State University of New York at Buffalo
Buffalo, NY 14260
U.S.A.

ABSTRACT

The similarity between the semantic network method of knowledge representation in artificial intelligence and shastric Sanskrit was recently pointed out by Briggs. As a step towards further research in this field, we give here an overview of semantic networks and natural-language understanding based on semantic networks. It is shown that linguistic case frames are necessary for semantic network processing and that Sanskrit provides such case frames. Finally, a Sanskrit-based semantic network representation is proposed as an interlingua for machine translation.

1 This material is based in part upon work supported by the National Science Foundation under Grant No. IST-8504713 (Rapaport) and in part by work supported by the Air Force Systems Command, Rome Air Development Center, Griffiss Air Force Base, NY 13441-5700, and the Air Force Office of Scientific Research, Bolling AFB, D.C. 20332 under contract No. F30602-85-C-0008 (Srihari).
1. INTRODUCTION

Computational linguistics is a subfield of artificial intelligence (AI) concerned with the development of methodologies and algorithms for processing natural-language by computer. Methodologies for computational linguistics are largely based on linguistic theories, both traditional and modern. Recently, there has been a proposal to utilize a traditional method, viz., the shastric Sanskrit method of analysis (Briggs, 1985), as a knowledge representation formalism for natural-language processing. The proposal is based on the perceived similarity between a commonly used method of knowledge representation in AI, viz., semantic networks, and the shastric Sanskrit method, which is remarkably unambiguous.

The influence of Sanskrit on traditional Western linguistics is acknowledgedly significant (Gelb, 1985). While linguistic traditions such as Mesopotamian, Chinese, Arabic, etc., are largely enmeshed with their particularities, Sanskrit has had at least three major influences. First, the unraveling of Indo-European languages in comparative linguistics is attributed to the discovery of Sanskrit by Western linguists. Second, Sanskrit provides a phonetic analysis method which is vastly superior to Western phonetic tradition and its discovery led to the systematic study of Western phonetics. Third, and most important to the present paper, the rules of analysis (e.g., sutras of Panini) for compound nouns, etc., is very similar to contemporary theories such as those based on semantic networks.

The purpose of this paper is threefold: (i) to describe propositional semantic networks as used in AI, as well as a software system for semantic network processing known as SNePS (Shapiro, 1979), (ii) to describe several case structures that have been proposed for natural-language processing and which are necessary for natural language understanding based on semantic networks, and (iii) to introduce a proposal for natural-language translation based on shastric Sanskrit and semantic networks as an interlingua.

2. KNOWLEDGE REPRESENTATION USING SEMANTIC NETWORKS

2.1. Semantic Networks

A semantic network is a method of knowledge representation that has associated with it procedures for representing information, for retrieving information from it, and for performing inference with it. There are at least two sorts of semantic networks in the AI literature (see Findler 1979 for a survey): The most common is what is known as an “inheritance hierarchy,” of which the most well-known is probably KL-ONE (cf. Brachman & Schmolze 1985). In an inheritance semantic network, nodes represent concepts, and arcs represent relations between them. For instance, a typical inheritance semantic network might represent the propositions that Socrates is human and that humans are mortal as in Figure 1(a). The interpreters for such systems allow properties to be “inherited,” so that the fact that Socrates is mortal does not also have to be stored at the Socrates-node. What is essential, however, is that the representation of a proposition (e.g., that Socrates is human) consists only of separate representations of the individuals (Socrates and the property of being human) linked by a relation arc (the “ISA” arc). That is, propositions are not themselves objects.

[Figure 1 here]

In a propositional semantic network, all information, including propositions, is represented by nodes. The benefit of representing propositions by nodes is that propositions about propositions can be represented with no limit. Thus, for example, the information represented in the inheritance network of Figure 1(a) could (though it need not) be represented as in Figure 1(b); the crucial difference is that the propositional network contains nodes (m3, m5) representing the propositions that Socrates is human and that humans are mortal, thus enabling representations of beliefs and rules about those propositions.
2.2. SNePS

SNePS, the Semantic Network Processing System, is a knowledge-representation and reasoning software system based on propositional semantic networks. It has been used to model a cognitive agent’s understanding of natural-language, in particular, English (Shapiro 1979; Maida & Shapiro 1982; Shapiro & Rapaport 1986, 1987; Rapaport 1986). SNePS is implemented in the LISP programming language and currently runs in Unix- and LISP-machine environments.

Arcs merely form the underlying syntactic structure of SNePS. This is embodied in the restriction that one cannot add an arc between two existing nodes. That would be tantamount to telling SNePS a proposition that is not represented as a node. Another restriction is the Uniqueness Principle: There is a one-to-one correspondence between nodes and represented concepts. This principle guarantees that nodes will be shared whenever possible and that nodes represent intensional objects. (Shapiro & Rapaport 1987.)

SNePS nodes that only have arcs pointing to them are considered to be unstructured or atomic. They include: (1) sensory nodes, which—when SNePS is being used to model a cognitive agent—represent interfaces with the external world (in the examples that follow, they represent utterances); (2) base nodes, which represent individual concepts and properties; and (3) variable nodes, which represent arbitrary individuals (Fine 1983) or arbitrary propositions.

Molecular nodes, which have arcs emanating from them, include: (1) structured individual nodes, which represent structured individual concepts or properties (i.e., concepts and properties represented in such a way that their internal structure is exhibited)—for an example, see Section 3, below; and (2) structured proposition nodes, which represent propositions; those with no incoming arcs represent beliefs of the system. (Note that structured proposition nodes can also be considered to be structured individuals.) Proposition nodes are either atomic (representing atomic propositions) or are rule nodes. Rule nodes represent deduction rules and are used for node-based deductive inference (Shapiro 1978; Shapiro & McKay 1980; McKay & Shapiro 1981; Shapiro, Martins, & McKay 1982). For each of the three categories of molecular nodes (structured individuals, atomic propositions, and rules), there are constant nodes of that category and pattern nodes of that category representing arbitrary entities of that category.

There are a few built-in arc labels, used mostly for rule nodes. Paths of arcs can be defined, allowing for path-based inference, including property inheritance within generalization hierarchies (Shapiro 1978, Sripahi 1981). All other arc labels are defined by the user, typically at the beginning of an interaction with SNePS. In fact, since most arcs are user-defined, users are obligated to provide a formal syntax and semantics for their SNePS networks. We provide some examples, below.

Syntax and Semantics of SNePS

In this section, we give the syntax and semantics of the nodes and arcs used in the interaction. (A fuller presentation, together with the rest of the conversation, is in Shapiro & Rapaport 1986, 1987.)

(Def. 1) A node dominates another node if there is a path of directed arcs from the first node to the second node.

(Def. 2) A pattern node is a node that dominates a variable node.

(Def. 3) An individual node is either a base node, a variable node, or a structured constant or pattern individual node.

(Def. 4) A proposition node is either a structured proposition node or an atomic variable node representing an arbitrary proposition.

(Syn.1) If \( w \) is a(n English) word and \( i \) is an identifier not previously used, then
NATURAL LANGUAGE PROCESSING

\[
\begin{array}{c}
\text{LEX} \\
\text{i} \rightarrow \text{w}
\end{array}
\]

is a network, \( w \) is a sensory node, and \( i \) is a structured individual node.

(Sem.1) \( i \) is the object of thought corresponding to the utterance of \( w \).

(Syn.2) If either \( t_1 \) and \( t_2 \) are identifiers not previously used, or \( t_1 \) is an identifier not previously used and \( t_2 \) is a temporal node, then

\[
\begin{array}{c}
\text{BEFORE} \\
\text{t}_1 \rightarrow \text{t}_2
\end{array}
\]

is a network and \( t_1 \) and \( t_2 \) are temporal nodes, i.e. individual nodes representing times.

(Sem.2) \( t_1 \) and \( t_2 \) are objects of thought corresponding to two times, the former occurring before the latter.

(Syn.3) If \( i \) and \( j \) are individual nodes and \( m \) is an identifier not previously used, then

\[
\begin{array}{c}
\text{PROPERTY} \\
\text{j} \rightarrow \text{m} \rightarrow \text{i}
\end{array}
\]

is a network and \( m \) is a structured proposition node.

(Sem.3) \( m \) is the object of thought corresponding to the proposition that \( i \) has the property \( j \).

(Syn.4) If \( i \) and \( j \) are individual nodes and \( m \) is an identifier not previously used, then

\[
\begin{array}{c}
\text{PROPER-NAME} \\
\text{j} \rightarrow \text{m} \rightarrow \text{i}
\end{array}
\]

is a network and \( m \) is a structured proposition node.

(Sem.4) \( m \) is the object of thought corresponding to the proposition that \( i \)'s proper name is \( j \). (\( j \) is the object of thought that is \( i \)'s proper name; its expression in English is represented by a node at the head of a LEX-arc emanating from \( j \).)

(Syn.5) If \( i \) and \( j \) are individual nodes and \( m \) is an identifier not previously used, then

\[
\begin{array}{c}
\text{CLASS} \\
\text{j} \rightarrow \text{m} \rightarrow \text{i}
\end{array}
\]

is a network and \( m \) is a structured proposition node.

(Sem.5) \( m \) is the object of thought corresponding to the proposition that \( i \) is a (member of class) \( j \).

(Syn.6) If \( i \) and \( j \) are individual nodes and \( m \) is an identifier not previously used, then

\[
\begin{array}{c}
\text{SUPERCLASS} \\
\text{j} \rightarrow \text{m} \rightarrow \text{i}
\end{array}
\]

is a network and \( m \) is a structured proposition node.

(Sem.6) \( m \) is the object of thought corresponding to the proposition that (the class of) \( i \)s are (a subclass of the class of) \( j \)s.

(Syn.7) If \( i_1 \), \( i_2 \), \( i_3 \) are individual nodes, \( t_1 \), \( t_2 \) are temporal nodes, and \( m \) is an identifier not previously used, then

[Figure 2 here]

is a network and \( m \) is a structured proposition node.
(Sem.7) \( m \) is the object of thought corresponding to the proposition that agent \( i_1 \) performs act \( i_2 \) with respect to \( i_3 \) starting at time \( t_1 \) and ending at time \( t_2 \), where \( t_1 \) is before \( t_2 \).

3. NATURAL-LANGUAGE UNDERSTANDING USING SEMANTIC NETWORKS

Semantic networks can be used for natural-language understanding as follows. The user inputs an English sentence to an augmented-transition-network (ATN) grammar (Woods, 1970, Shapiro 1982). The parsing component of the grammar updates a previously existing knowledge base containing semantic networks (or builds a new knowledge base, if there was none before) to represent the system's understanding of the input sentence. Note that this is semantic analysis, not syntactic parsing. The newly built node representing the proposition (or a previously existing node, if the input sentence repeated information already stored in the knowledge base) is then passed to the generation component of the ATN grammar, which generates an English sentence expressing the proposition in the context of the knowledge base. It should be noted that there is a single ATN parsing-generating grammar; the generation of an English output sentence from a node is actually a process of "parsing" the node into English. If the input sentence expresses a question, information-retrieval and inferencing packages are used to find or deduce an answer to the question. The node representing the answer is then passed to the generation grammar and expressed in English.

Here is a sample conversation with the SNePS system, together with the networks that are built as a result. User input is on lines with the : prompt; the system's output is on the lines that follow. Comments are enclosed in brackets.

: Young Lucy petted a yellow dog
I understand that young Lucy petted a yellow dog

[The system is told something, which it now "believes." Its entire belief structure consists of nodes b1, m1-m13, and the corresponding sensory nodes (Figure 3). The node labeled "now" represents the current time, so the petting is clearly represented as being in the past. The system's response is "I understand that" concatenated with its English description of the proposition just entered.]

: What is yellow
a dog is yellow

[This response shows that the system actually has some beliefs; it did not just parrot back the above sentence. The knowledge base is not updated, however.]

: Dogs are animals
I understand that dogs are animals

[The system is told a small section of a class hierarchy; this information does update the knowledge base.]

[Figure 3 here]

There are three points to note about the use of SNePS for natural-language understanding. First, the system can "understand" an English sentence and express its understanding; this is illustrated by the first part of the conversation above. Second, the system can answer questions about what it understands; this is illustrated by the second part. Third, the system can incorporate new information into its knowledge base; this is illustrated by the third part.

Case Frames

Implicit in such a language understanding system are so-called case frames. We give a brief summary here; for a more thorough treatment see Winograd (1983).
Case-based deep structure analysis of English was suggested by Fillmore (1968). The surface structure of English relies only on the order of constituents and propositions in a clause to indicate role. Examples are:

Your dog just bit my mother.
My mother just bit your dog.

In Russian, Sanskrit, etc., explicit markings are used to represent relationships between participants. Examples in Russian, which uses six cases (nominative, genitive, dative, accusative, instrumental, and prepositional), are:

Professor uchenika tseloval (the professor kissed the student).
Professora uchenik tseloval (the student kissed the professor).

The extremely limited surface case system of English led Fillmore to suggest cases for English deep structure as follows: Agentive (animate instigator of action), Instrumental (inanimate force or object involved), Dative (animate being affected by action), Factitive (object resulting from action), Locative (location or orientation), and Objective (everything else). For example, consider the sentence:

John opened the door with the key.

Its case analysis yields: Agentive = John, Objective = the door, Instrumental = the key.

Schank (1975) developed a representation for meaning (conceptual dependency) based on language independent conceptual relationships between objects and actions: case roles filled by objects (actor, object, attribuant, recipient), case roles filled by conceptualizations (instrument, attribute, ...), and case roles filled by other conceptual categories (time, location, state). For example:

John handed Mary a book.

has the analysis: Actor = John, Donor = John, Recipient = Mary, Object = book, Instrument = an action of physical motion with actor = John and object = hand.

4. Sanskrit Case Frames and Semantic Networks

In the previous section we noted that natural-language understanding based on semantic networks involves determining what case frames will be used. The current set of case frames used in SNePS is not intended to be a complete set. Thus, we propose here that shastric Sanskrit case frames, implemented as SNePS networks, make an ideal knowledge-representation "language."

There are two distinct advantages to the use of classical Sanskrit analysis techniques. First, and of greatest importance, it is not an ad hoc method. As Briggs (1985) has observed, Sanskrit grammarians have developed a thorough system of semantic analysis. Why should researchers in knowledge representation and natural-language understanding reinvent the wheel? (cf. Rapaport 1986). Thus, we propose the use of case frames based on Sanskrit grammatical analysis in place of (or, in some cases, in addition to) the case frames used in current SNePS natural-language research.

Second, and implicit in the first advantage, Sanskrit grammatical analyses are easily implementable in SNePS. This should not be surprising. The Sanskrit analyses are case-based analyses, similar, for example, to those of Fillmore (1968). Propositional semantic networks such as SNePS are based on such analyses and, thus, are highly suitable symbolisms for implementing them.

As an example, consider the analysis of the following English translation of a Sanskrit sentence (from Briggs 1985):

Out of friendship, Maitra cooks rice for Devadatta in a pot over a fire.

Briggs offers the following set of "triples," that is, a linear representation of a semantic network for this sentence (Briggs 1985: 37, 38):

cause, event, friendship
friendship, object1, Devadatta
NA TURAL LANGUAGE PROCESSING

friendship, object2, Maitra
cause, result, cook
cook, agent, Maitra
cook, recipient, Maitra
cook, instrument, fire
cook, object, rice
cook, on-loc, pot

But what is the syntax and semantics of this knowledge-representation scheme? It appears to be rather ad hoc. Of course, Briggs only introduces it in order to compare it with the Sanskrit grammatical analysis, so let us concentrate on that, instead. Again using triples, this is:

cook, agent, Maitra
cook, object, rice
cook, instrument, fire
cook, recipient, Devadatta
cook, because-of, friendship
friendship, Maitra, Devadatta
cook, locality, pot

Notice that all but the penultimate triple begins with cook. The triple beginning with friendship can be thought of as a structured individual: the friendship between Maitra and Devadatta. Implemented in SNePS, this becomes the network shown in Figure 4. Node m11 represents the structured individual consisting of the relation of friendship holding between Maitra and Devadatta. Node m13 represents the proposition that an agent (named Maitra) performs an act (cooking) directed to an object (rice), using an instrument (fire), for a recipient (named Devadatta), at a locality (a pot), out of a cause (the friendship between the agent and the recipient).

Such an analysis can, presumably, be algorithmically derived from a Sanskrit sentence and can be algorithmically transformed back into a Sanskrit sentence. Since an English sentence, for instance, can also presumably be analyzed in this way (at the very least, sentences of Indo-European languages should be easily analyzable in this fashion), we have the basis for an interlingual machine-translation system grounded in a well-established semantic theory.

[Figure 4 here]

5. INTERLINGUAL MACHINE TRANSLATION

The possibility of translating natural-language texts using an intermediate common language was suggested by Warren Weaver (1949). Translation using a common language (an “interlingua”) is a two-stage process: from source language to an interlingua, and from the interlingua to the target language (Figure 5). This approach is characteristic of a system in which representation of the “meaning” of the source-language input is intended to be independent of any language, and in which this same representation is used to synthesize the target-language output. In an alternative approach (the “transfer” approach), the results of source text analysis are converted into a corresponding representation for target text, which is then used for output. Figure 6 shows how the interlingua (indirect) approach compares to other (direct and transfer) approaches to machine translation (MT). The interlingua approach to translation was heavily influenced by formal linguistic theories (Hutchins 1982). This calls for an interlingua to be formal, language-independent, and “adequate” for knowledge representation.

[Figures 5 and 6 here]
NATURAL LANGUAGE PROCESSING

Various proposals for interlinguas have included the use of formalized natural-language, artificial "international" languages like Esperanto, and various symbolic representations. Most prior work on interlinguas has centered on the representation of the lexical content of text. Bennet et al. (1986) point out that a large portion of syntactic structures, even when reduced to "canonical form," remain too language-specific to act as an interlingua representation. Thus, major disadvantages of an interlingua-based system result from the practical difficulty of actually defining a language-free interlingua representation.

Besides, none of the existing MT systems use a significant amount of semantic information (Slocum 1985). Thus, the success of an interlingua depends on the nature of the interlingua as well as the analysis rendered on the source text to obtain its interlingual representation. This made the interlingua approach too ambitious, and researchers have inclined more towards a transfer approach.

It has been argued that analyses of natural-language sentences in semantic networks and in Sanskrit grammar is remarkably similar (Briggs 1985). Thus we propose an implementation of Sanskrit in a semantic network to be used as an interlingua for MT. As an interlingua, Sanskrit fulfills the basic requirements of being formal, language-independent, and a powerful medium for representing meaning.

REFERENCES


Figure 1(a). An “ISA” inheritance-hierarchy semantic network

Figure 1(b). A SNePS propositional semantic network (m3 and m5 represent the propositions that Socrates is human and that humans are mortal, respectively)
Figure 2.
Figure 3. SNePS network for "Young Lucy petted a yellow dog."
Figure 4. SNePS network for "Out of friendship, Maitra cooks rice for Devadatta in a pot over a fire."
Figure 5. The Interlingua Approaches
Figure 6. Various Approaches to MT
TECHNICAL REPORTS

Following are recent technical reports published by the Department of Computer Science, University at Buffalo. Reports may be ordered from:

Library Committee
Department of Computer Science
University at Buffalo (SUNY)
226 Bell Hall
Buffalo, NY 14260
USA

Prices are given as (USA and Canada/Other Countries). Payment should be enclosed with your order (unless we have a free exchange agreement with your institution, in which case the reports are free). Please make checks payable to “Department of Computer Science”.

If your institution is interested in an exchange agreement, please write to the above address for information.

86-01 A. Rastogi, S.N. Srihari: Recognizing Textual Blocks in Document Images Using the Hough Transform
($1.00/$1.50)

($1.00/$1.50)

86-03 W.J. Rapaport: Philosophy of Artificial Intelligence: A Course Outline
($1.00/$1.50)

($1.00/$1.50)

86-05 S.L. Hardt, W.J. Rapaport: Recent and Current AI Research in the Dept. of Computer Science, SUNY-Buffalo
($1.00/$1.50)

86-06 P.J. Eberlein: On the Schur Decomposition of a Matrix for Parallel Computation
($1.00/$1.50)

86-07 R. Bharadhwaj: Proto-Maya: A Distributed System and its Implementation
($2.00/$3.00)

86-08 J. Geller, J.J. Hull: Open House 1986: Current Research by Graduate Students of the Computer Science Dept., SUNY-Buffalo
($1.00/$1.50)

86-09 K. Ebcioglu: An Expert System for Harmonization of Chorales in the Style of J.S. Bach
($3.00/$4.00)

86-10 S.C. Shapiro: Symmetric Relations, Intensional Individuals, and Variable Binding
($1.00/$1.50)

86-11 S.N. Srihari, J.J. Hull, P.W. Palumbo, C. Wang: Automatic Address Block Location: Analysis of Images and Statistical Data
($1.00/$1.50)

($1.00/$1.50)

86-13 G.L. Sicherman, J.Y. Arrasjid: The Buffalo Font Catalogue, Version 4
($1.00/$1.50)

86-14 J.M. Wiebe, W.J. Rapaport: Representing De Re and De Dicto Belief Reports in Discourse and Narrative
($1.00/$1.50)

($1.00/$1.50)
<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>86-16</td>
<td>P.J. Eberlein</td>
<td>Comments on Some Parallel Jacobi Orderings</td>
<td>$1.00/$1.50</td>
</tr>
<tr>
<td>86-17</td>
<td>P.J. Eberlein</td>
<td>On One-Sided Jacobi Methods for Parallel Computation</td>
<td>$1.00/$1.50</td>
</tr>
<tr>
<td>86-18</td>
<td>R. Miller, Q.F. Stout</td>
<td>Mesh Computer Algorithms for Computational Geometry</td>
<td>$1.00/$1.50</td>
</tr>
<tr>
<td>86-19</td>
<td>E.J.M. Morgado</td>
<td>Semantic Networks as Abstract Data Types</td>
<td>$3.00/$4.00</td>
</tr>
<tr>
<td>86-20</td>
<td>G.A. Bruder, J.F. Duchan, W.J. Rapaport, E.M. Segal, S.C. Shapiro, D.A. Zubin</td>
<td>Deictic Centers in Narrative: An Interdisciplinary Cognitive-Science Project</td>
<td>$1.00/$1.50</td>
</tr>
<tr>
<td>86-21</td>
<td>Y.H. Jang</td>
<td>Semantics of Recursive Procedures with Parameters and Allasing</td>
<td>$1.00/$1.50</td>
</tr>
<tr>
<td>86-22</td>
<td>D.S. Martin</td>
<td>Modified Kleene Recursion Theorem Control Structures</td>
<td>$1.00/$1.50</td>
</tr>
<tr>
<td>86-23</td>
<td>P.J. Eberlein</td>
<td>On using the Jacobi Method on the Hypercube</td>
<td>$1.00/$1.50</td>
</tr>
<tr>
<td>86-24</td>
<td>W.J. Rapaport</td>
<td>Syntactic Semantics: Foundations of Computational Natural-Language Understanding</td>
<td>$1.00/$1.50</td>
</tr>
<tr>
<td>87-01</td>
<td>S. Chakravarty, H.B. Hunt III</td>
<td>On Computing Test Vectors for Multiple Faults in Gate Level Combinational Circuits</td>
<td>$1.00/$1.50</td>
</tr>
<tr>
<td>87-02</td>
<td>J. Case</td>
<td>Turing Machines</td>
<td>$1.00/$1.50</td>
</tr>
<tr>
<td>87-03</td>
<td>S.N. Srihari, W.J. Rapaport, D. Kumar</td>
<td>On Knowledge Representation using Semantic Networks and Sanskrit</td>
<td>$1.00/$1.50</td>
</tr>
</tbody>
</table>