Computer Science 4150
Artificial Intelligence
Winter, 2005

Instructor:
Nick Cercone - CSBldg 105
902-494-2832 - nick@cs.dal.ca

Semantic Networks

Note: This is a foundational article, newer applications of semantic networks and references are available but they do not advance the notational efficacy nor expressive power of the material presented in this article.

1.0 Introduction

Semantic networks have stimulated debate about their use as a propositional knowledge representation in reasoning and understanding systems since their introduction by Quillian (1968, 1969) and subsequent generalization by other researchers including Shapiro (1971, 1979, 1987), Schubert (1975, 1976), Woods (1975), Levesque & Mylopoulos (1979), etc. Early on, semantic networks were successfully exploited for concept learning (Winston, 1970), natural language understanding (Schank, 1972, 1973; Schubert et al., 1979) and deductive reasoning (McSkimmon & Minker, 1977). They were also influential in psychological theories of cognition (Collins & Quillian, 1972; Norman & Rumelhart, 1975; and Wilson, 1979), the early development of knowledge representation languages (Brachman, 1979; Levesque & Mylopoulos, 1977), the implementation of semantic network processing systems (Shapiro, 1979), and machine architectures (Fahlman, 1975, 1979, 1982). More recently, variants of semantic networks have been utilized in various inheritance mechanisms.

Initial misconceptions about the use of semantic networks in knowledge representation were widespread. By the end of the 1970's, however, both semantic network and predicate calculus representations were recognized as formal languages designed to allow natural language statements to be paraphrased precisely and unambiguously and whose respective reputations could be attributed to the use to which each had been put. Research efforts based on semantic networks emphasized associative and other non-deductive processing whereas the predicate calculus was usually wedded to resolution-based theorem proving. It then became widely recognized that theorem proving techniques could just as easily be adapted to semantic network representations or non-deductive inference algorithms could be designed to predicate calculus.

The natural development and influence of semantic networks can be found in current knowledge representation and object-oriented languages. Data engineers and knowledge representation researchers, concerned with the need to formally understand the expressive power and representational adequacy of various data and knowledge representation schemes, have returned to logic to provide a basis for such reflections. “Recasting” various formalisms in logical terms provides a basis for comparison and has led to the development of new non-monotonic logics to handle incompleteness, default reasoning and other unique requirements. These developments are having an effect on semantic network proponents; McCalla & Cercone (1983) and Cercone & McCalla (1987) provide good overviews of network-based and other data and knowledge representations.

In this paper we report on what has come to be known as the ECO family formalism of semantic network. After briefly reviewing the early and ad-hoc development of semantic networks, we evolve the formalism due to Schubert (1975) and its continued development in over a decade's subsequent work in the representation and organization of knowledge, accelerated special purpose inference mechanisms designed around this graphical two dimensional logic representation and its use in an English COnversational System.

2.0 The Ad-Hoc Development of Semantic Nets

Early artificial intelligence (AI) representations of knowledge were developed in an ad-hoc manner, largely in response to the constraints of implementation. This is particularly true of semantic network representations of knowledge. A varied assortment of semantic networks have appeared in many AI systems as a means for representing knowledge. They have been used to represent the meanings of English words, as representations of static information (facts) for deductive planning systems and question-answering programs, and as crude knowledge organizational schemes for both frame and non-frame formalisms. Until the late 1970's, semantic networks had been used in informal and disparate ways that have precluded their precise, nonrestrictive definition.

Quillian's Initiation

Quillian (1968) developed what came to be known as the semantic network as the result of his pioneering effort to model semantic memory and explain the organization of semantic information in human memory. In Quillian's networks, word meanings were represented by a network of objects and relations among the objects. To distinguish ambiguous meanings for pairs of words Quillian compared the plausibilities of different interpretations by the strength of the associations linking the pairs of meanings. Quillian distinguished the ***type node*** whose associative links lead directly into a configuration of other nodes that comprise the meaning of the object represented by the type node and the ***token node*** which refers indirectly to a word concept by having one special kind of associative link that points to the concept's type node. Figure 2-1 illustrates word meanings for the three meanings of plant: (1) a living structure which is not an animal, frequently with leaves, getting its food from air, water, earth; (2) apparatus used for any process in industry; and (3) put (seed, plant, etc.) in earth for growth. The three circled words, plant1, plant2, and plant3 represent type nodes and all other words represent token nodes. The non-terminated links from token nodes represent pointers to the token's type node.

Quillian's effort initiated a surfeit of network formalisms from which several interesting associative processing algorithms developed, for example Winston's network matching algorithm, Schank et al.'s language processing heuristics, etc. A unique network representation was characterized by the kinds of nodes, associative links, and types of operations that could be carried out within their network framework. Without exception, these networks were expressively weak.

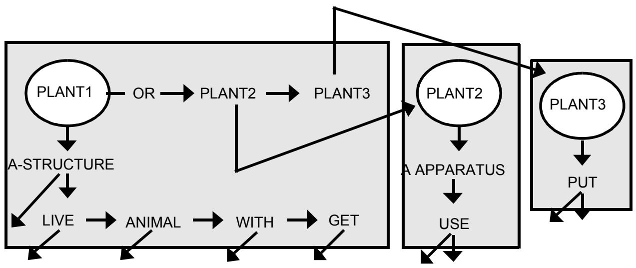


Figure 2-1. Three meanings of plant.

Winston's Semantic Network

When presented with line drawings of scenes containing children's toy blocks, such as bricks, cubes, pyramids and wedges, Winston's (1970) program forms descriptive networks for these scenes disclosing the property and relationships of the objects appearing in them. Using these structural descriptions, the program can learn structural concepts such as “pedestal”, “arch”, or “arcade” on the basis of examples and counterexamples of the concepts.[1]

After determining the bodies in a scene, the program then determines which edges belong to which object, and fills in partially occluded edges. Then the program infers the types of objects (brick, wedge, etc.) from the shapes and adjacency relationships of the visible faces. The sizes and orientation are then readily available. The example illustrated in Figure 2-2a is in the form of a semantic network. Nodes depict particular things (such as the object ABC, and its faces A, B, and C) or general concepts (such as BRICK, LARGE, etc.), and the arcs are relations between things and/or general concepts (e.g., PART-IS is a relation which holds between a thing and its parts).

[1] While the content of the scenes is severely restricted, the methods employed in the program seem generalizable to more utilitarian concept-formation tasks (consider “table”, “chair”, etc.). Winston-type structural learning evolved significantly from learning from parameter adjustment which had been the major dominant paradigm in pattern recognition up to that point. The work of Havens & Mackworth (1983, 1987) represents a major deviation in recognition tasks and is one of the dominant paradigms for scene recognition at the present time.

Heuristic routines are next applied to scenes to obtain support and relative position relationships between bodies. This information is represented in descriptive network form, see Figure 2-2b. The final scene description includes these relationships as well as the overall attributes of the object determined earlier (brick, wedge, etc.) but not the finer details such as component faces and their shapes.

Winston's system learns using network descriptions and a clever network matching algorithm is developed. For example, to teach the concept house, the system is presented an instance and three non-instances of houses, Figure 2-3a. The initial model formed by the system describes the first true instance of the concept. The model is generalized for subsequent examples so that it will accept any new instances and reject non-instances.[2] Each modification of the model is made by generating a comparison network for the current model and the new example. A comparison network describes the similarities and differences between the model and the example. The description of scenes (1) and (2), their comparison network, and subsequent modifications of the model are shown for the “house” sequence in Figure 2-3b.

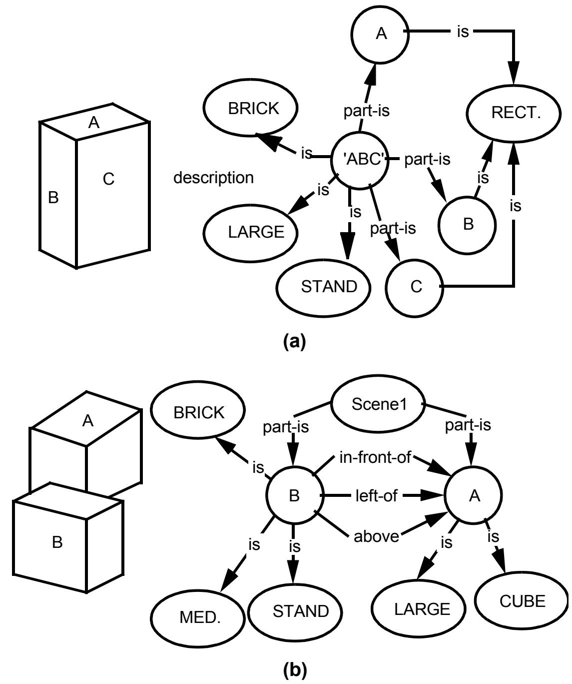


Figure 2-2. (a) A large standing brick (b) A scene description.

Note the absence of the SUPPORTS pointer in the second network. In the comparison network there is an extra pointer in the left network, labeled SUPPORTS with destinations the node for the wedge. The modified model asserts there is necessarily a support relation between the two parts of the scene. This modified model will reject (2) as an instance of a house. Similarly (3) cause a reinforcement of the

[2] Non-instances are required to be fairly close to the true instances so that the system does not have too difficult a time determining which features of a non-instance disqualify it.

wedge property of the supported object by the necessary operator, and (4) causes a reinforcement of the brick property of the supporting object. Just as certain features of the model can be reinforced through counter-examples, others can be relaxed when true instances are presented which differ from the model.

Although Winston's program embodies many ad-hoc decisions, both in analyzing scenes and in the learning process and contains logical inadequacies in the network formalism (disjunction and quantification cannot be represented), the program incorporates a semantic network formalism and representation and a non-trivial algorithm for comparison networks.

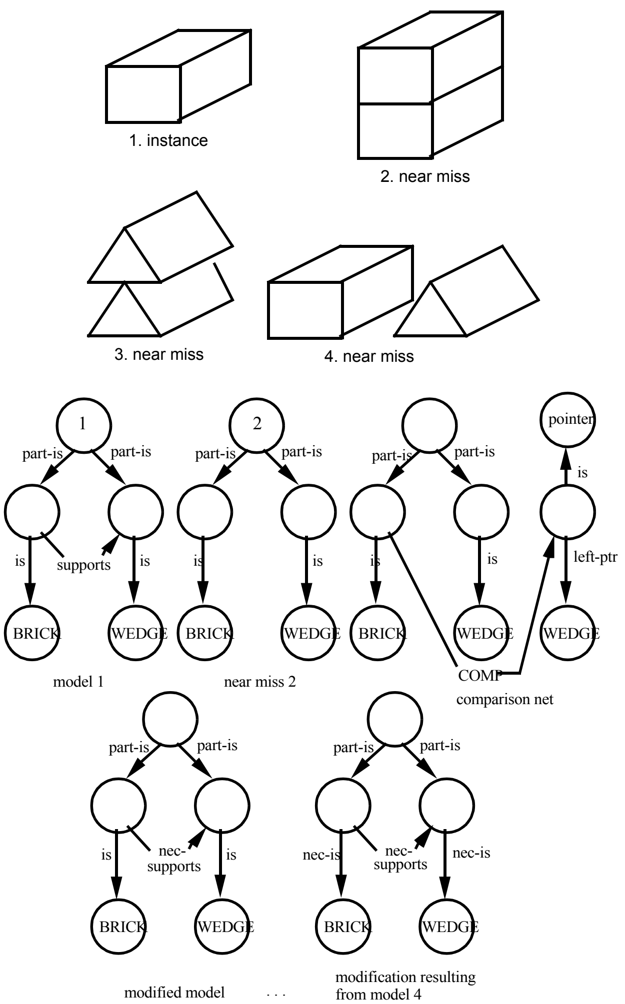


Figure 2-3. (a) A house and three near misses. (b) Description of scenes and comparison network.

Rumelhart, Lindsay, and Norman's Process Model for L-T Memory

Rumelhart, Lindsay, and Norman (1972) made one of the first attempts to formalize a semantic network in their work on modelling long-term memory. Their active structural network was a labelled connected graph consisting of a set of nodes interconnected by a set of relations associating two nodes. They (arbitrarily) distinguished primary nodes, which refer directly to a natural language concept, and secondary nodes, which represent a concept as it is used in a specific context - a token (in Quillians' sense).

Rumelhart et al. gave formal definitions for their notation and rules for forming relations, concepts (qualifiers, quantifiers, propositions, etc.), propositions, and operators. Application of their rules yield Figure 2-4 as an encoding of the sentence “John and Mary want (to have) three red balloons”. This representation incorporated case structures into a network formalism. Figure 2-4 typifies general definitions of concepts incorporated into the network in a straightforward manner with case-like pointers indicating parts of nominal concepts and agents and objects of verbs. They used the infamous ISA link to indicate type-token relations as well as subset relations; other links were poorly motivated and left unexplained. Little attention was paid to the logical adequacy of the representation thus the relationship between the diagrams and the concepts which they represent are left unclear.

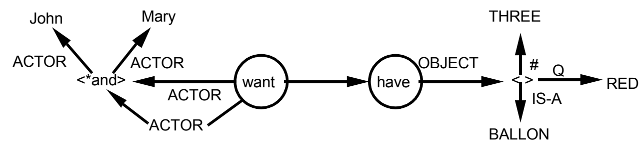


Figure 2-4. John and Mary want three red balloons.

Rumelhart's group did attempt to account for procedural information in their representation via the ISWHEN link and they integrated case information with other aspects of world knowledge. The LUIGI system described in Scragg’s Ph.D. thesis (1975), answers questions about processes and makes use of the model and representation developed by Rumelhart et al. Inconsistencies tend to diminish the appeal of their model. Their introduction of binary predicates isa and has property is a trick used by advocates of semantic networks to artificially convert unary to binary predicates. This is logically unnecessary unless one's purpose is to convert a higher-order logic into a many-sorted first-order logic.

A serious impediment is the lack of distinction between general and specific (individual) concepts. Other difficulties centre around adverbial modification. They treat adverbial modifiers as operators that apply to relations and other operators to generate new relations or operators. Unfortunately many adverbial modifiers require systematic analysis rather than mere replacement by n-valued relations.

Simmons Computational Representation and Use

Simmons (1973) formally defined semantic nets according to the rules

network	:=	node* {* means one or more repetitions, the modified Kleene *}
node	:=	atom + relationset | terminal constant
atom	:=	Ci | Li {indexed contextual meaning or lexical meaning}
relationset	:=	relation + node
relation	:=	member of a list of semantic relations
terminal constant	:=	character string {examples are “noun”, “sing”, “act”}

Simmons then explained how to compute semantic networks from English strings using a variant of Woods' (1970) augmented transition network (ATN) grammars and the generation of English sentences from the semantic network. Simmons used this formulation of a semantic network in a relatively simple question-answering system. Nevertheless, the expressive power and representation adequacy of networks was not a concern of Simmons.

What are Semantic Networks and Why Use Them

As described, semantic networks have been used in different applications and have been specified by notational variants. Shapiro (1971) most succinctly stated the distinctive characteristic of semantic networks, "All the information about a given conceptual entity should be reachable from a common place". Thus it is that semantic nets can easily be thought of as clever indexing schemes for propositional knowledge representations, the network concepts represent intuitively meaningful entities and knowledge about these entities is directly attached to them.

The succinctness, clarity, and intuitive nature of semantic networks argues in their favour if only for purely methodological advantages. Semantic networks are readable; they suggest procedures for comprehension and inference, and the computer data structures which they resemble. The examples given demonstrate how associative processing algorithms and complex pattern matching operations were readily identifiable using networks. Although in both Quillian's and Winston's applications, the semantic nets each utilized were weak expressively, it is doubtful that either would have developed his particular associative processing algorithms without the use and perspicuity of semantic networks.

3.0 Extending the Expressive Adequacy of Semantic Networks: A Basic Network Formalism

We examine semantic network representations of knowledge, discussing their suitability as a representation of propositional knowledge. The use of semantic networks as a uniform representation mediating between specialized representations appropriate to particular task domains is considered.

3.1 A Data Structure for Semantic Information Processing

Early on, Shapiro (1971, 1979) attempted to extend the expressive power of semantic networks. He introduced the important distinction between the conceptual relation and the structural relation. Structural relations are used to form the basic structure of the semantic network represented exclusively by non-conceptual arc labels whereas conceptual relations are represented by nodes. Shapiro also calls structural relations item relations and conceptual relations are referred to as system relations.

Shapiro's repertoire of meaningful entities included individuals (particular concepts), properties and relations (generic concepts), and propositions. The knowledge attached to a concept is the set of propositions in which it participates plus, possibly, relevant procedures.

A network syntax allowing arbitrary nesting of quantifiers and propositional operator scopes is essential and one such syntax is discussed in the next section. The first complete representation for quantifiers was introduced by Shapiro. He treated every and some as relations between sentences and individuals (the variables are assumed quantified) occurring in those sentences. Thus a quantified sentence such as "Everyone loves Mary" involves two propositions nodes: one for the open sentence [[x person]?[x loves Mary]] and another for the proposition that the relation every holds between x and the open sentence. This method of quantification is syntactically complete but seems semantically unsatisfying since unbound variable nodes, open sentence nodes, and relations over such nodes are not intuitively meaningful.

The use of Skolem functions to represent quantification in networks dispenses with variable-binding operators and thus introduces no meaningless nodes. For example, $\forall$x $\exists$y [y taller-than x] becomes [y(x) taller-than x]. The universal quantification of x is implicit, and y(x) is the Skolem function supplying a specific individual that is "taller-than" x corresponding to each x.

The importance of Shapiro's early contribution was largely ignored until the mid 1970's when the concern for foundations in knowledge representation theory became of paramount importance.

3.2 A Comprehensive Network Formalism

A comprehensive network formalism is motivated and explained in Schubert et al. (1979). Special problems with respect to the use of logical connectives, quantifiers, descriptions, modalities, and some other constructions that fail in conventional semantic networks are systematically resolved with extensions to conventional network notations. The representation harmonizes with linear one dimensional logical notations, illustrating the close kinship of the two notations. This kinship supports the claim that networks have inherited formal interpretability from logical notations.

To unify previous network formalisms and extend their expressive power to include quantifiers, operators, and higher-order predication, Schubert et al. developed a canonical form of a semantic network. Schubert's network notation is a uniform representation, mediating between the special purpose representations which are necessary for representing and processing different task domains. It is intended as a standard of comparison and serves to illustrate concepts of language comprehension.

We highlight the basic notation developed earlier, and present a few more complex examples to illustrate the *expressive power, notational efficacy* and *logical adequacy* of the network formalism.

3.2.1 The Basic Notation

Semantic networks are graphical analogues of data structures that represent facts in a computer system. *Solid loops* are used as nodes that represent either explicitly diagrammed proposition nodes or existentially quantified concept nodes. In the first case, propositions form the basic unit of knowledge represented by the notation. In the later case, nodes may be labeled with names for the concepts they denote, for example, John, book, book1, ...; ordinary attributive terms such as book are reserved for the corresponding universal concepts, while numerically suffixed words such as book1 are used for particular instances of the concepts. *Unbroken lines* are used as arcs linking parts of a proposition to proposition nodes. *Arc labels* are simply distinguishing marks. They are analogous to parentheses or commas in the predicate calculus in that they relate denoting terms syntactically; they are non-denotative themselves. Whenever possible they will be chosen to enhance readability and be suggestive of meaning, but numeric labels could also be used, cf McDermott (1975). To avoid confusion, predicate names will be designated in small letters and arc labels by capital letters. An example of a simple network is offered in Figure 3-1a. An English paraphrase of Figure 3-1a is: proposition P1 represents the English sentence “John loves Mary”, P2 represents “Helen loves John”, P3 represents “Helen dislikes Mary”, P4 represents “John gave Mary a red dress”, P5 and P6 serve to identify the particular red dress which John gave to Mary.

Occasionally the detailed use of arcs and nodes in the explicit notation of Figure 3-1a will clutter a diagram, reducing readability. Figure 3-1b illustrates an abbreviated form of 3-1a with the understanding that the structure is built upon explicit propositions. The full network is abbreviated by coalescing the predicate node into the proposition node, removing the solid loop.

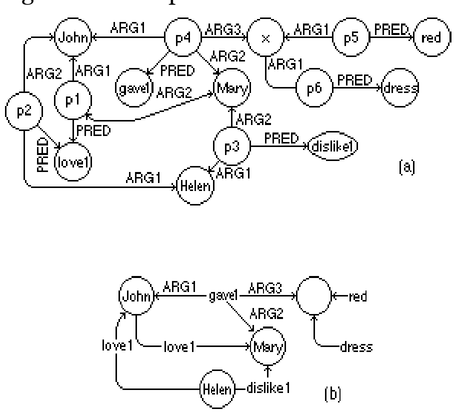


Figure 3-1. (a) A Simple Network (b) its abbreviated form.

3.2.2 A Comprehensive Network Formalism

The basic semantic network was extended with notation for logical connectives, conventions for indicating quantifier and operator scopes, including n-ary and higher-order predicates, and providing formal interpretability for each construction. The notation extends the expressive power of the network, making it equivalent to high-order and modal logics.

Notation for Logical Connectives

Logical connectives, ignored in most network formalisms, occur frequently in discourse and are necessary for truth-functional completeness. In Schubert's semantic network notation logical connectives are represented as explicit nodes for logical compounds of propositions (or open sentences), with graphical links to the components. Figure 3-2 represents the English sentence “If either the Mets or the Pirates win and the Expos take second place, then I'll recover past losses and either buy a colour TV or fly to Greece”. The figure illustrates the formation of disjunctions and conjunctions explicitly by the use of graphical links to tokens of the disjunction and conjunction operators. The operator-operand links of the logical operator are represented by *broken lines* in Figure 3-2. Observe that no distinguishing marks are needed on the links of disjunction or conjunction (they are symmetrical operators) and arrowheads can be dropped when there is no ambiguity. The use of past as a modifier of losses is an evasive manoeuvre; it postpones discussion of time. Other logical connectives can be introduced in the same way.

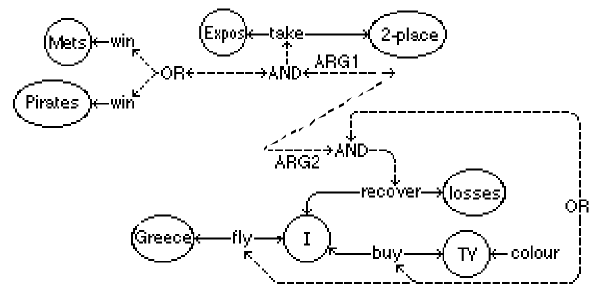


Figure 3-2. If either the Mets or Pirates win and the Expos take second place, then I'll recover my past losses and either buy a colour TV or fly to Greece.

In a semantic network containing logical compounds it is not suitable to regard all propositions in the network as asserted. In this formalism we adopt the convention that the complete semantic net asserts exactly those propositions which are not constituents of compound propositions (that is, operands of connectives or modal operators). Graphically this means that precisely those propositions are asserted which are not pointed to. We must devise a method for asserting a proposition which is also a constituent of a compound proposition. The assertion of a constituent simplifies any logical compound. For propositional attitudes, causes, intentions and the like, however, it may prove worthwhile to assert a proposition independently of the compound. In this case, we can use disjunction with a single operand, V ? p, as a way of saying p holds since the compound proposition established by the token V is not pointed to.

Representing Quantifiers in the Semantic Network

One prominent use of semantic network representations includes natural language understanding systems. Any representation of natural language must include quantifiers such as “All boy scouts do good deeds”, and “several of my friends were on strike today”. General knowledge relies upon quantifiers as in “It is always raining on Long Beach”. Definite descriptions implicitly make use of quantification as the example “the people of China” shows. The meanings of complex action concepts require quantification, such as “walking” which has associated with its definition, assertions such as at all times, some of the limbs of the individual engaged in walking support the individual.

Quantification challenges semantic network representations to indicate the scopes of universal and existential quantifiers. The notation used is analogous to quantifier-free normal form in predicate calculus. Propositions are in prenex form (quantifiers have maximum scope), existentially quantified variables are Skolemized, and universal quantification is implicit. To distinguish between them, we simply use *solid loops* for existentially quantified concept nodes (as in all previous figures), and *broken loops* for universally quantified nodes. Graphical Signalization then links each existentially quantified node to all universally quantified nodes on which it depends (that is, whose universal quantifiers precede the existential quantifier in prenex form). *Dotted lines* represent these dependency links to distinguish them from propositional and logical links. For example, "Every critic admires some painting" is represented as shown in Figure 3-3a. In predicate calculus notation this is (∀x)[critic(x) ? (∃y) (painting(y) & admires(x,y))] or "critic(x)? [painting(f(x)) & admires (x,f(x))]", Skolemized. Now if we can assume (∃y)(painting(y), that is, there is at least one painting (or alternatively, that there is at least one critic), then this becomes painting(f(x)) & [critic(x) => admires (x,f(x))] which corresponds to the slightly simpler diagram shown in Figure 3-3b. Here the painting proposition is no longer considered as a consequent of the critic proposition. This type of simplification is often suitable for encoding natural language statements, since we do not usually communicate in terms of propositions which are trivially true by virtue of the nonexistence of their referents (which is not to say that we do not communicate about nonexistent entities). The diagram for "There is a painting which all critics admire" differs from Figure 3-3b only in the absence of the dependency link between the painting and critic nodes.

The proposed method of representing quantification is applicable only to propositions in prenex form. If we deal only with existential logic (one in which all propositional constructions are truth-functional), no generality is lost, though clarity is occasionally compromised. However, propositions involving (nonextensional) modal operators such as necessarily and believes cannot be converted to prenex form. To represent such propositions, the present notation is expanded to allow arbitrary embedding of quantifiers. The generalized scope notation allowing non-prenex propositions is illustrated in Figure 3-3c. Scope inclusion links establishing operator precedence over quantifiers run from proposition nodes to variables, not from operator nodes to variables. If the diagram were given in full form rather than the abbreviated notation this would be explicit. Other examples of higher-ordered constructions are in Schubert et al. (1979).

Techniques for representing knowing and believing apply equally to other propositional attitudes such as remembering, supposing, intending, deciding, avoiding, hoping, imagining, pretending, and trying. Nonreferential terms within the scopes of such operators (whichever ones are deemed useful independently of the others) can be identified by means of scope dependency links. The same applies to the denotic modalities such as obligation. It should be obvious, for example, how "John ought to marry the prettiest girl" would be represented.

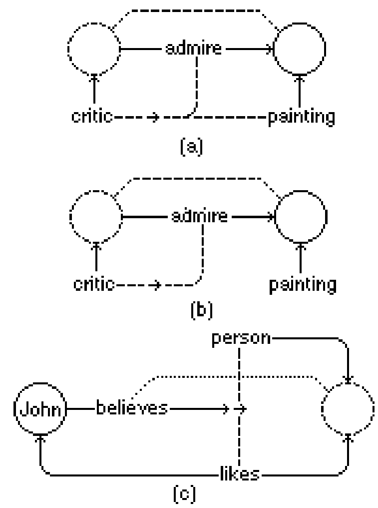


Figure 3-3. (a&b) Every critic admires some painting. (c) John believes that everyone likes him.

Causal dependency is an important modality deserving attention. "John asked Mary to dance because she was the only girl left without a partner" is an example of an opaque context generated by a causal construction. Substitution of the term Mary for its referential synonym the only girl left without a partner clearly fails. As in other modal constructions, we may need scope dependency links to express causal structures.

Logical quantifiers are unsuitable for expressing many natural language quantifiers such as several, many, most of, a few, more than, etc. These quantifiers can be handled systematically with (fuzzy) properties of set cardinality and relations between set cardinalities, plus standard set relations such as set inclusion. We can classify natural language quantifiers according to their indication of set size from absolute to comparative, where comparative indicators are those comparing the size of one set to that of another set. The logical quantifier ∃ (there exists) is an absolute indicator of set size since (∃x)P(x) tells us that the set of P's contains at least one member. The quantifier ∀ (for all), by virtue of its equivalence to ~∃~, is also an absolute indicator. In the context (∀x)[P(x)?Q(x)] however, where the number of P's is finite, it can be considered a comparative indicator of set size. It tells us that the subsets of P's that are Q's is as large as the set of P's itself. Common absolute quantifiers are none, one, two, three, ..., several; common comparative quantifiers are all of, most of, a small fraction of, a slight majority of, one-half of, two-thirds of, as many as, twice as many as, etc. Some quantifiers show both absolute and comparative attributes, especially some and many. For example, in "many artificial satellites are orbiting the globe" *many* is used absolutely - it appears to imply a cardinality of at least about a dozen. In "many students attend John's class" *many* is used in the sense considerably more than attend the average class. This particular use of many is discussed in Bartsch & Vennemann (1972). They do not consider the absolute indicativeness of many, however, nor of its comparative use in selecting a subset of another set, as in "many of the world's people are undernourished". Contrast the numerical indication here with that in "many of the apples in the basket were rotten".

Recognizing the absolute/comparative behaviour of quantifiers, we can describe them systematically by means of predicates on set cardinality and on pairs of set cardinalities. In Figure 3-4 the convention for abbreviating implication is shown, that is, *single broken lines* are used for the conjoined antecedents and *solid lines* for the conjoined consequences. The predicate # denotes the number of elements of a set, or rather the single-valued relation which holds between a set and the number of its elements. We regard # as a function from sets onto integers and several as a (fuzzy) property of numbers. If we accepted "several" as a possible value of set size, then it would be impossible to talk about the size of the set, as # would be many valued (for example, a 6-element set might have both size 6 and size *several*).

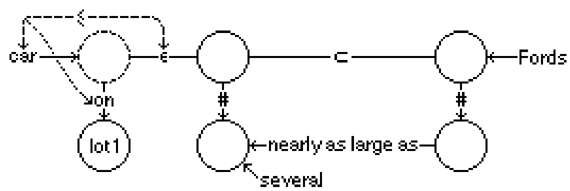


Figure 3-4. Several cars were on the lot; most of them were Fords.

Representing Definite and Relational Descriptions

Natural language representations usually ignore the distinction between definite and indefinite descriptions, so the method for representing both types of descriptions is shown here. The representation of both definite descriptions (the little old lady at the door) and indefinite descriptions (a big apple) are based on the conventions for logical connectives and quantifiers already introduced rather than on a description operator such as Moore (1973) uses. Description operators are useful only at a superficial level of language representation, and in the domain of pure mathematics.

The description "John's car" conveys the presupposition that John has exactly one car. This is true at least for certain discourse contexts. The non-equivalent proposition that John owns a red car does not necessarily convey the uniqueness condition has been expressed in John's car. The use of equality leads to a convention which is suitable for definite descriptions of sets, such as "the SOCRED supporters of British Columbia", as well as for "John's car". Figure 3-5 illustrates the representation of a definite description of sets.

The notation for descriptions as introduced is inadequate for descriptions of predicative concepts which are expressed in terms of predicates of the same, rather than higher, type. For example, suppose we wish to say that the property human is the same as the property rational animal; note that the latter property is of type 1 and is expressed in terms of the type 1 properties rational and animal. We cannot diagram this statement on the basis of the formula (∀x)[human(x)?rational(x) & animal(x)], since this merely asserts extensional identity (that is, the set of human beings equals the set of rational animals). The desired statement of intensional identity can be made with the aid of Church's lambda (λ) operator. This operator abstracts a predicate from an open sentence by designating certain variables of the sentence as arguments of the predicate. Thus we write human = λx [rational(x) & animal(x)]. Two additional interesting examples are provided by the sentences "Loving one's neighbors is a virtue", which requires abstraction of the monadic predicate loves one's neighbors from the dyadic predicate loves, and "The sun is shining brightly on the beach" which requires formation of the λ-abstracted sun shining brightly on the beach for comparison with the sun which is shining brightly on the beach. This latter example is shown in Figure 3-6, using a graphical analogue of λ-abstraction.

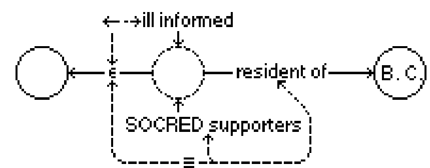


Figure 3-5. The SOCRED supporters of B. C. are illinformed.

(∃S)[(∀x)[member(x,S)?[SOCRED-supporters(x) & resident-of(x,B.C.)]] & (∀x)[member(x,S)?illinformed(x)]]

λ-conversion is accomplished by means of a (solid) *λ-link* from the proposition expressing that an individual loves all of his neighbours to the node for the individual. In general graphical λ-abstraction involves the construction of some open sentence (possibly with embedded λ-expressions), and λ-conversion of some of the variables of the sentence. Open sentences are formed exactly like propositions, except that some of the participating concept nodes are regarded as free variables. λ-conversion is symbolized by λ links from the node corresponding to an open sentence to free variables of the sentence. In the nonmonadic case λ links are labeled λ-A, λ-B, etc., (or in some other systematic way) to distinguish the arguments of the abstracted predicate.

Including Time in the Representation

Time is regarded as the only situational or contextual variable, cf. McCarthy & Hayes (1969) that needs to be added to action propositions. This contrasts with Anderson & Bower (1973), Schank et al. (1975), etc. who include time and locale as basic dimensions of events in their representation. However, locale is not a property of events as a whole, but a (frequently time-dependent) property of the participants in an event. For example, in "John is watching a circling hawk" it is John and the hawk who have locations, not the event.

We use time in ***instantaneous*** (a proposition can have either a fixed or variable moment of time associated with it) and ***interval*** modes. In the interval mode the moment is replaced by a time interval which can be omitted in particular contexts to simplify propositions that describe states and events with more enduring properties (like being a girl, car, etc.). This omission is a matter of expediency; any change involving a metamorphosis (a caterpillar becoming a butterfly) would require explicit recognition of time dependencies.

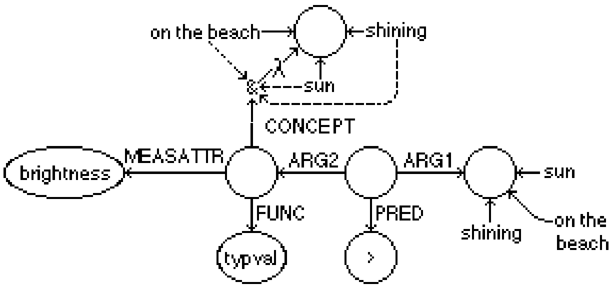


Figure 3-6. The sun is shining brightly on the beach.

Within this framework, temporal relations including tenses (which can be built up from more elementary temporal relations) can be defined. If we restrict our view of time as consisting of a set of elements (time points) and a relation that partially orders them, we can define binary temporal relations similar to those of Bruce (1972). Bruce devised a systematic method for defining tenses. He mapped time relations given by auxiliary verbs and the form of the main verb and defined seven binary ordering relations on time segments, for example, I had gone - maps to after. Thus a tense is an n-ary relation on time segments, for example, past tense is one in which the relation after holds between two time segments. Our modifications to some of Bruce's binary ordering relations permit us to represent a sentence like "While he was in Rome, before he met his murderer, he first sang in La Gravity" as in Figure 3-7.

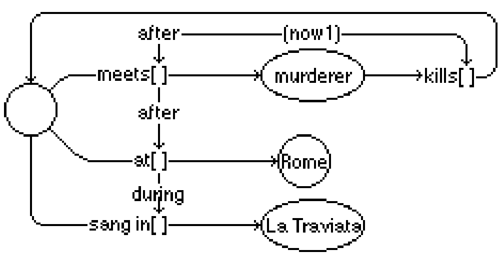


Figure 3-7. While he was in Rome, before he met his murderer, he first sang in La Traviata.

Primary issues of network form and content need to be addressed, including: (i) primitive versus nonprimitive representations; (ii) the separation of propositional content of text from pragmatic aspects; and (iii) network normal form versus ad hoc systems. Computer systems design for specific tasks depends in part on early commitments to these issues. These ideas are detailed elsewhere, see Schubert et al. (1979), Brachman (1979), Woods (1975), etc., and summarized next.

4.0 What Should be Represented, Why, and How

Application designers face fundamental representation, organization, and processing issues early in their approach to design, and choices are critical. The designer must decide what should be represented, the form of the representation, and the level to which the representation is restricted. On the basis of these considerations, we argue in favour of a non-primitive semantic network representation in which propositions are organized in normal form determined by the concept hierarchy.

4.1 The Problem with Semantic Primitives

The expressive power and formal interpretability of the *state-based representation* (Cercone and Schubert, 1975) contrasts with Schank's (1972) *conceptual dependency* representation and Wilks' (1973) *preference semantics* representation. We compare the methods of Schank and Wilks with the network oriented state-based representation in order to argue against the use of a small number of very general primitive predicates for representing meaning in natural language.

Conceptual dependency epitomized in the MARGIE system, Schank et al. (1973), is rich in semantic representation and designed to assist paraphrase, inference, and machine translation. Schank represents meaning structures with a graphical notation that divides words into four conceptual categories. In addition, Schank uses only 14 primitive actions from which all other actions are derived.

Wilks' preference semantics also utilizes primitives to represent the meaning content of natural language statements. Wilks concentrated on machine translation of small input paragraphs and reported reasonably good translation from English to French. Wilks' representation is based almost entirely on about 60 primitive predicates. Furthermore, Wilks (1977) admonishes Hayes (1974, 1977) for presenting a number of arguments against the use of semantic primitives as Schank and Wilks use them.[1]

Wilks misconstrues Hayes' remarks when he ascribes to Hayes the belief that a coherent and consistent metaphysics for STUFF is necessary for all ordinary language comprehension. At the other extreme, embedding the minimal content of terms into a minimum conceptualization does not facilitate the human interpretive process. The original term itself suggests what content we could infer in addition to the minimal content. This idea of inference can be efficiently programmed in a semantic structure by inserting probable inferences with direct reference to the word definitions. This is simpler than analyzing the minimal representation and then looking for applicable inference rules.

---

1 Wilks writes "One aspect of these criticisms is not radical - in the sense of questioning the very basis of primitives - but it is a demand by Hayes that primitive systems give a more explicit account of the rules regulating inferences concerning a primitive for substance, like STUFF. This demand for greater explicitness is a good one, though there is reason to doubt that any coherent and consistent metaphysics of substance can in fact be given. Two and a half millennia of philosophy have failed to provide one, yet throughout that time everyday conversation about substances, such as coal, oil, and air goes on unimpeded. It is important to stress this fact, so as not to fall into the error of imagining that language about substances requires such a metaphysics of substances in order to function at all. It clearly does not."

Wilks rejects Hayes' criticism that there is no model theoretic semantic for primitive based systems. He feels that Hayes' demand for such a model theoretic semantics makes Hayes' demand for a metaphysics of STUFF radical. Wilks emphatically rejects the application of model theoretic semantics (in the manner of the semantics that Tarski constructed for logic) to the analysis of natural language meaning. Wilks believes that preference semantics evolve inevitably into a natural language itself. However, Wilks misconstrues truth conditions as serving to determine the actual truth of sentences in the object language, and gives the example of the inappropriateness of computing over a possible world. However, possible worlds are not intended to be computational domains, but as part of an abstract conception of meaning and truth. Truth is thus only relevant to truth-determination. Model theoretic semantics does, provide a practical means to decide truth-determination, e.g., checking whether an inference mechanism is truth-preserving.

Wilks also chides Bobrow (1975) for arguing that a primitive expansion or paraphrase requires a more complex match than does the original English word that the paraphrase is for. He disputes the complexity of the matching, however, since preference semantics does not operate in paraphrase mode, he uses Schank's arguments about the paraphrase mode of Schank's primitive based system to reject Bobrow's critique. Examining Schank's defense of primitive-based systems, we find the following advantages: (1) paraphrase relations are made clearer; (2) similarity relations are made clearer; (3) inferences that are true of various classes of verbs can be treated as coming from the individual (primitive) ACTs. The inferences come from ACTs and states rather than from words; and (4) organization in memory is simplified because much information need not be duplicated. The primitive ACTs provide focal points under which information is organized.

The increased clarity of paraphrase and similarity relations derives from Schank's use of canonical form rather than his “primitives” meaning representation, see Schubert et al. (1979) for detailed arguments. The last two advantages also cannot be traced to the use of semantic primitives. Sharing inferences within classes of verbs can be accomplished without restating words in terms of primitive ACTS. Cercone (1975) gives an example which demonstrates both eats and drinks as sentential forms share in the implications that a single primitive ingests would store but conserve storage and computation.

Moreover, while we see no disadvantages of non-primitive based representations, point (4) shows a major disadvantage in their elimination, namely the resultant need for matching complex primitive representations instead of originally simple propositions. Examining a typical restaurant script such as Schank proposes for John dined at a restaurant convinces us of the complexity of matching. Schank's method stores John's actions in the restaurant as a sequence of scenes, partially obtained from the restaurant script, which represent several successive conceptualizations about restaurant dinings. An inquiry such as “Did John dine at a restaurant?” requires another construction of the complex succession of conceptualizations about restaurant dining. Then the succession of conceptualizations would have to be matched. The task is not trivial. Wilks entire primitive-based system is spared this complication since the system was designed for sentence translation and not question answering.

Schubert et al. (1979) present arguments which clarify the need for a meaning representation more detailed than Schank and Wilks' oversimplified meaning formulas. We maintain that no representation of meaning is adequate until it captures many of the same notions that people realize when they comprehend language utterances. Our version of what walking means to people is at least an order of magnitude more complex than the definitions Schank and Wilks allow for walking, since their formulas also admit other complex concepts such as running, skipping, skating, and hopping under the same meaning formula. Our example demonstrates that the semantic network formalism was equal to the representation task and also dramatized the actual complexity of ordinary concepts when expressed in the primitive-based representations of Schank and Wilks. The meaning formulas of primitive-based representations highlight properties most frequently needed for comprehension and simple inferences. This is their remaining salient feature. Primitive-based representations do capture major properties of defined concepts and we have only added minor details to them. But to rely on meaning caricatures as Schank and Wilks do, ensures that comprehension will remain of a crude sort. Non-primitive based representations can be equipped with the advantages of the Schank-Wilks approach, simply by providing lists of the most frequently needed properties for comprehension of each predicate and permitting the significant properties of concepts to become independently accessible without invoking the full meaning representation defining the concept.

The complexity of a concept does not interfere with its matchability since it is retrieved by its name. Considerations of storage economy and the computational complexity of pattern-directed retrieval convince us of the limited value of primitive-based representations.

4.2 Propositional Content and Pragmatic Aspects

Is it sufficient to extract the propositional content of sentences and use that as the basis of representation, or should the representation also reflect aspects of meaning such as speaker intention, presupposition, connotation, and style? An example from Woods (1975) demonstrates the significance of these various aspects of meaning. Woods argues that the sentence

The dog that bit the man had rabies. [4.1]

should not have a representation identical to the sentence

The dog that had rabies bit the man.[4.2]

even though the propositional content for [4.1] and [4.2] is identical, that is, there is a dog, there is a man, the dog had rabies, and the dog bit the man (ignoring, for simplicity, temporal considerations). Woods protests, correctly, that the differing descriptions of [4.1] and [4.2] are inappropriate criteria for accessing the memory node for the referent of either description. Woods further insists that [4.1] and [4.2] intuitively mean different things, so syntactic distinctions must be made between the meaning expressed in the relative clauses and the meaning expressed in the main clauses. Perhaps Woods derives this position because he believes that intensional and extensional entities must be represented by different sorts of nodes in a semantic network. For example, Woods says that in some contexts the prettiest blonde refers to only Sally Sunshine, yet in other contexts the prettiest blonde depends on the notion conveyed by the descriptive phrase. Woods believes that these contexts are distinguished by different sorts of nodes (or sub-networks). We believe that terms (or nodes) already encompass both extensions and intentions, and that a syntactic distinction is not appropriate to distinguish extensional and intensional nodes. It is appropriate to explain the conditions under which a term contributes to the truth value of a sentence through its intension rather than through its extension alone. Woods' differentiation between intensional and extensional entities parallels the distinction between transparent and opaque readings which can be illustrated by the sentence “John wants to marry the prettiest girl”. The syntactic distinction lies in the relative scope of the *wants* modal operator and the existential quantifier of the example sentence.[1]

We propose that a distinction be made between the propositional content of sentences and their pragmatic aspects. Different pragmatic aspects generate the different meanings of sentences like [4.1] and [4.2]. We agree with Woods that the internal meaning representation of a sentence should reflect both its propositional content and its pragmatic aspects, but the two sorts of information should not be inextricably mixed. Mingling propositional and pragmatic information would handicap comprehension processes which must utilize any acquired knowledge. Woods' special syntactic representational device would also encumber the matching process since the matching processes seeking suitable referents for [4.1] and [4.2] would depend on the original text. In contrast, Schank (1972) has presented convincing reasons why an internal representation should be in a canonical form, relatively independent of the original English sentence. Mingling propositional and pragmatic information about utterances would disperse pragmatic information about a particular section of discourse over the propositional data base. Information about speaker intentions and assumptions would be buried with knowledge about dogs, people, etc. We maintain that a separate model for discourse status (speaker intentions and the like) is necessary. This model is the proper place for semantic information.

4.3 Network Form

We limit our discussion of the form of representation to the issue of property inheritance. In a separate paper, Vogel, Cercone, and Popowich (1990) examine property inheritance much more exhaustively using a new paradigm of beliefs and socially determined context.

The extended semantic network notations are capable of expressing any arbitrary proposition expressible in English, for example Schubert et al (1979), Levesque et al. (1979). But any system designed for reasoning about the real world must also effectively exploit property inheritance within generalization hierarchies. Conceptual entities typically consist of many components, the relationship between these components is valuable information. We require a mechanism which allows inheritance of the relationships from components to corresponding components within a conceptual entity. For example, the attachment relationships between the body parts of birds would require nontrivial inference processes to transfer to other similarly structured animals.

The method of variable-sharing was proposed by Hayes (1977) and adapted by Schubert et al. (1979) to solve this problem and allow for trivial transfer of relationships. We recommend that the knowledge associated with a generalization hierarchy be stored as a set of implicative propositions which share one universally quantified node and any number of existentially quantified nodes dependent on the universally quantified node. The antecedents of the implications involve the universally quantified node as argument and correspond to concepts making up the generalization hierarchy in the manner of hierarchies described above. Thus the implicants of a concept are accessible by topic rather than a long list of propositions involved in the concept. This mechanism facilitates addition of new information and we speculate that it is possible to organise other than monadic concepts, say relational concepts, hierarchically as well.

5.0 Using Semantic Networks for Knowledge Representation

A general theory of natural language understanding requires representations with sufficient expressive power to represent the meaning content of ordinary language. Schubert's extended semantic network notation possesses this expressive power. We now present a development of some ideas concerning the semantic network representation of individual items of factual knowledge in a computer, where this knowledge is thought of as being conveyed to the computer in natural language.

5.1 The Basic Framework

The basic framework embodied by many natural language understanding systems is the <actor-action- object> formalism. Certainly this is not without justification. Much of natural thought and communication follows from this framework. Only a deeper analysis of actions and intentions seems to belie thinking of this framework as underlying natural language. While not denying the intrinsic value to organization, heuristic programming, and pragmatics that this <actor-action-object> formalism suggests, its theoretic value for beginning investigations into language comprehension is minimal.

[1] For sake of simplicity we ignore the additional ambiguity of whether the prettiest girl refers to the time of John's wanting or the time of John's marrying.

A more fundamental starting point uses extended semantic networks to represent many natural language constructions in a neutral state-based representation, explicitly representing the propositional content of language utterances. At times, explicit comparisons are made between conceptual dependency theory and preference semantics theory on the one hand and the extended network approach on the other.

5.2 States, Events, Actions, Cases, Causes, and Intentions

Sentences such as

> **The sun was turning red and approaching the western horizon.[5.1]**

raises many questions about Schank's and Wilks' formalisms. In [5.1] the motion of the sun *must* be done by somebody or something whereas its change of colour *cannot* be done by somebody or something. Thus, using the <actor-action-object> formalism espoused by both Schank and Wilks, modes of behaviour which are expressed by actions *must* have actors whereas all other modes of behaviour cannot have actors. In the case of the (apparently) moving sun in sentence [5.1], one is hard pressed to identify the actor. Consequently we are compelled to regard certain ongoing activities which intuitively just happen as instigated by someone or something (including natural forces in a vague, unspecified sense).

Just as we are compelled to regard certain ongoing activities as instigated by somebody or something, we are denied the option of regarding certain actions as having an agent as in

> **John was hurting Mary by pulling her hair.[5.2]**

In [5.2] the hurting not being an action, has no actor whereas in

> **John was dragging Mary by pulling her hair. [5.3]**

the dragging, insofar as it involves PTRANS'ing does have John as an actor.

We may wonder by what criterion we draw the line between what an actor does and what he causes. In [5.2], according to Schank, we are to regard the "hurting" as caused by the "pulling" action. But the same is true of PTRANS'ing in [5.3]. Furthermore, even direct bodily action such as moving an arm can be viewed as caused by muscle contraction or, subjectively, as caused by an act of will, either of which again may have antecedent causes.

It seems to us that no structural primitives should be associated with actors at all. Instead we propose a neutral representation in which events are expressed as sequences of states of the participants. The successive states simply express what happened, without explicit commitment as to who did it. Agent(s) in an event can be identified by supplementary propositions. The notion of an agent can continue to be used to aid interpretation and inference. Agent would be regarded as a rather fuzzy higher level concept, understood by the system in terms of the role of a supposed agent within a sequence of causally and teleologically related states. In the sentence "John uprooted the sapling" the term agent can be considered highly applicable to John's role in the event while in the sentence "The avalanche uprooted the tree" its applicability to the role of the avalanche would be considered relatively low. The notion of an agent seems to depend in part on causal priority of a state of the supposed agent in the sequence of states under consideration, and in part on the extent to which purposive behaviour can be ascribed to the supposed agent in general, and in part to the extent to which the particular sequence of states which he initiated can be assumed to be intentional on his part.

Similarly we propose to separate *why* something happened (causes, enabling conditions, reasons, explanations, justifications, and the like) from *what* happened. This does not prevent us from including causal propositions in the representation and relying heavily on them for interpretation and inference. However, time relations and changes of state, not causes, will give coherence to a set of propositions as an event.

Schank's instrumental case relation between actions can and should be represented in terms of causation and intention. If a system has a conceptualization to the effect that John was PTRANS'ing the ball by PROPEL'ing it, then this conceptualization should also express that the PROPEL'ing was causing the PTRANS'ing. Phrases ostensibly expressing instrumental actions often express no more than causation. An example is the by clause in

> **The effluents were killing the fish by raising the temperature of the water.[5.4]**

When there is a difference, it lies in the intimation of purposive causation. In

> **John woke Mary by blowing his trumpet.[5.5]**

purposive causation is expressed, while in

> **Mary woke up because John was blowing his trumpet. [5.6]**

it is not. Sentences [5.5] and [5.6] clearly show that the instrumental relation amounts to a causal relation supplemented by intentional states.

Since the inception of conceptual dependency theory and preference semantics, there have been many changes to them. An early criticism of Cercone and Schubert (1975), apparently still somewhat valid, has largely been amended. Both theories now appear to have the conception of a state.

According to Fodor (1972), actions are to be thought of as a proper subclass of events. According to Schank, an action is something a nominal can be said to be doing at some moment.[1] A study of his proposed inferences shows that an action does not express a definite change in a situation; rather it expresses existence of a situation which tends to produce change, and all actual changes must be inferred. Formulas for actions in Wilks' theory are analogous but less explicit. Actions, then, express modes of behaviour which promote but do not guarantee the occurrence of events. For example, the

[1] this is an interpretation of Schank's definition.

actions PTRANS, INGEST, MOVE do not express changes in location; instead those changes are primary inferences given that an actor is PTRANS'ing, INGEST'ing, or MOVE'ing something. Syntactically, the relationship between an event, say a change in location, and the action, say PTRANS, whose primary inference is that event, corresponds quite closely to the relationship between verbs and their participles respectively. To say that John was PTRANS'ing himself with the result that his location changed is quite analogous to saying that he was going somewhere with the result that he went there.[1] In any case the term action is now seen to be quite misleading, since it normally connotes the occurrence of definite events, rather than the existence of a dynamic situation which tends to generate events.

Thus Schank's actions (contrary to the connotation of the term) correspond more closely to states than to events! To say that A is PTRANS'ing B is merely to express a momentary truth about the system in which A and B participate, not a change in that system (which remains to be inferred). This view is compatible with the observation that many common modifiers express subtle blends of passive and dynamic attributes. The examples below bring to mind conceptual images that illustrate a gradually increasing emphasis on dynamics.

blue sky ? *bright sun* ? *glowing (or luminous) candle* ? *burning candle* ? *blazing fire* ? *billowing smoke*

Schank's actions, and, as far as we can determine, Wilks', are dynamic states, or activities, or modes of behaviour which mediate changes in certain attributes. Thus PTRANS and MOVE mediate changes in location, INGEST and EXPEL mediate changes in containment, and MTRANS mediates changes in awareness.

We believe that the recognition that actions in Schank's sense are essentially states rather than events is important, since it leads to a uniform view of all (true) events as sequences of states. In this view the need for identifying actors of events does not arise, nor is it necessary to delineate the spurious boundary between passive and dynamic states.[2]

We now illustrate our representation of states and events. Nothing new needs to be added to the network notation of section 3. We regard any condition which can hold momentarily (blue, moving, running, etc.) as a state. Accordingly, any atomic proposition which is based on a time-dependent predicate is a state proposition. Figure 5-1 shows two concurrent state propositions: something (the redness of the sun) was increasing throughout some time interval and something else (the distance between the sun and the horizon) was decreasing throughout the same time interval.

Actually there are two additional state propositions, concerned with the existence of unique values of redness and distance at all moments of time within the time interval of interest; these have not been made explicit since they can be taken to be implicit in the redness and distance relations.

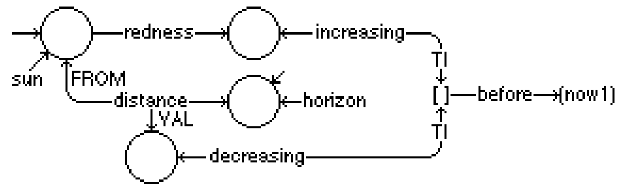


Fig 5-1. The sun was getting redder and approaching the horizon.

Actually there are two additional state propositions, concerned with the existence of unique values of redness and distance at all moments of time within the time interval of interest; these have not been made explicit since they can be taken to be implicit in the redness and distance relations.

Events involve a change in state as "the last leaf fell from the tree" illustrates. The definitive characteristic of state changes is the following: if a system has property A at time t1, and property B at time t2, then A?B is a change of state if and only if A and B are mutually exclusive properties, for example, A=solid, B=liquid; A=round, B=rectangular. In fact a state attribute such as colour which can assume various values can consistently be defined as a set of mutually exclusive properties, each member of the set being regarded as a value of the attribute. This admits both qualitative attributes such as colour as well as quantitative attributes such as location. Figure 5-2 shows a simple event involving a single change of state of a system with one component (Mary). The time relation then implies immediate succession of the two time intervals.

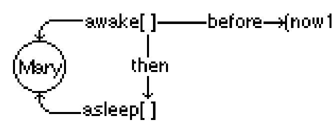


Figure 5-2. Mary fell asleep.

Our representation of Schank's sentence "John hit Mary" is shown in Figure 5-3a. A paraphrase is the following. Some unknown mode of behaviour of John caused some object to move quickly toward Mary. Subsequently the object reached Mary and exerted a force on her. Note that we have a state and an event here, viz. John's unknown state and the event of the object moving toward Mary and striking her. In accordance with our earlier remarks about causation, the causal connections between John's state and the ensuing event does not make John's state part of that event. Only exclusive and successive states of a particular system of objects form events. A natural inference in Figure 5-3a would be that John intentionally hit Mary, that is, that the missing state of John is that he was trying to bring about the event in question. We represent *trying* by the state predicate *x has active goal y at time t,* as illustrated in Figure 5-3b. The explanatory paraphrase goes as follows. The cow was on the ground, propelling itself towards a location above the moon; then it was

---

[1] Unlike Schank we do not regard *he was going* and *he went* as equivalent; we claim that *he went there*, unlike *he was going there*, affirms that he did arrive at his destination, **and** that it is decidedly odd to say *he went there but didn't get there.*

[2] Since the inception of conceptual dependency theory and preference semantics, there have been many changes to them. An early criticism of Cercone and Schubert (1975), apparently still somewhat valid, has largely been amended. Both theories now appear to have a conception of a state.

moving toward that location; then it was at that location; then it was moving towards a place of destination on the ground, such that the moon is between the place of departure and place of destination; then it was at the place of destination. Note that moving towards could have been represented in terms of distance decreasing as shown in Fig 5-1.

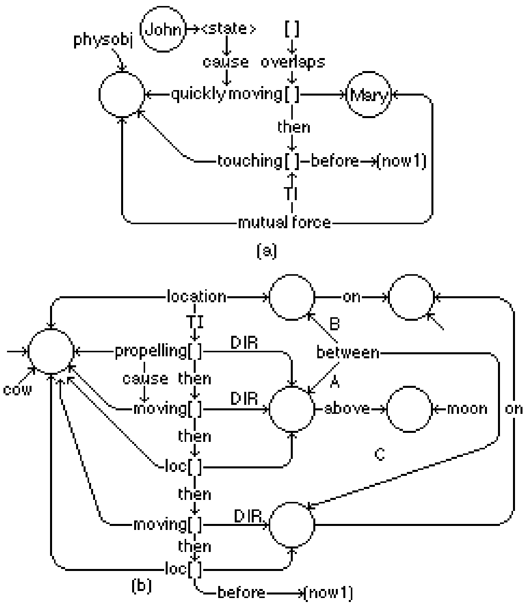


Fig 5-3. (a) John hit Mary. (b) The cow jumped over the moon.

An important consequence of our very broad conception of states is that new complex states (modes of behaviour) can be defined in terms of events involving primitive or previously defined states. The time of occurrence of these events can extend some distance backward and forward from the moment at which the new state is defined to hold. For example walking is defined in terms of successive states of motion and displacement of the walker's feet and body over a period of observation encompassing (say) two steps, since an instantaneous snapshot of a person is insufficient for deciding whether or not that person is walking (although it may of course supply enough cues to prompt the inference that the person is walking). A tentative definition of walking is given below.

Complex dynamic states (modes of behaviour) such as walking, running, dancing, tumbling, flickering, etc., can be constructed in terms of more elementary states. The constructions are necessarily as complex as the states they describe. Complexity can result from the intricate coordination of several simultaneous activities (for example, rolling expresses rotation and translation at coordinated rates), or from complex time dependencies (for example, flickering), or from both (for example, walking).

Since we intend to exploit fully the semantic "preferences" that any given predicate induces on its arguments, we propose to make use of *cases* in our approach to representation. For example the predicate *moving* prefers a physical object as its first argument and a physical location as its second argument; the predicate *has-active-goal* prefers a sentient being as its first argument and a state proposition as its second argument. Furthermore, there are broad similarities between the argument preferences of different predicates. For example, several predicates prefer animate objects in certain argument positions. We certainly can and sometimes do acknowledge such similarities and give a rough indication of the sort of preferences involved by using suggestive argument markers such as ANIM, THING, DIRECTION, PLACE, etc., instead of noncommittal markers such as A, B, C, ... , . However, we do not think that these markers can be chosen so that they express not merely similar but identical argument roles and semantic preferences, no matter in which predicate they occur.[1]

Thus semantic cases, while certainly useful heuristically in finding or inferring arguments of predicates have no universal or primitive status.

5.3 Complex Concepts

According to Schank's dictionary, if a human, X, walks to a location, Z, then X PTRANS's X by X MOVEing the feet of X in the direction of Z. This formula rules out walking on one's hands and knees, or walking on one's hands (admittedly a rare skill). More importantly, the formula admits running, skipping, hopping, jogging, shuffling, and even skating. Presumably, then, the dictionary entry is not intended to capture the full meaning of walking as we seem to understand it, but only those aspects which are most essential to language understanding and inference. Similarly Wilks' formulas are incomplete. For example, it is correct to say that DRINK implies ((*ANI SUBJ)(((FLOW STUFF) OBJE)((*ANI IN)(((THIS (*ANI (THRU PART))) TO) (BE CAUSE))))) but not the converse (which could mean someone was receiving an enema). So again a selection of only some linguistically important features has apparently been made.

It is important to formulate more complete meaning representations for two reasons. First, more information will be required for adequate comprehension of *ordinary* discourse. Second, much more information will surely be re-

[1] This view is supported by Bartsch and Vennemann (1972): "... case is entirely a surface category and not, as Fillmore (1968) suggests, a category of universal semantics. Semantic representations are based on propositions, which consist of a relation (n-ary predicate with n>=0) with a finite number of arguments filled either with constants or with bound variables. The "meaning" of an argument as argument is entirely determined by its relation. Therefore, no two arguments have precisely the same meaning, as arguments. Thus, if the meaning of an argument as argument is called a case, then there are as many cases as there are arguments, and this number, if it is finite at all, is a very large one. What some linguists call "cases" are classes of arguments based on certain semantic similarities which follow from the semantic similarities of their relations. The fact that certain arguments show similarities in their syntactic behaviour, such as tending to occur in certain position relative to the verb or belonging to the same surface case, does not support the assumption that there exists a small number of universal cases. Those syntactic similarities are simply a consequence of the fact that the human mind is structured in such a way that it tends to group objects on the basis of certain relevant similarities and then manipulate the objects of the group alike.

quired to match the human ability to describe concepts and reason about them. For example, suppose we ask a reasonably articulate person to describe human walking in as much detail as possible. We might elicit at least the following information: Each foot of the walker repeatedly leaves the ground, moves freely in the walking direction for a distance comparable to the length of the walker's legs (while staying close to the ground), then is set down again, and remains in position on the ground, supporting the walker, while the other foot goes through a similar motion. The repetition rate is about one repetition per second. The legs remain more or less extended. The body remains more or less erect and is carried forward at a fairly constant rate. Further details could be added about flexing motions of feet, knees, and hips, the slight up-and-down motion of the body, typical arm motion, and forces exerted on the ground. Figure 5-4 shows a network which describes walking (regarded as a state predicate with three arguments besides time) along these lines. A few propositions have been omitted so as not to clutter the diagram. These are that each foot is also above the ground (and close to it) while moving, that each foot is also supporting X while stationary; that the duration of each of the unlabeled time intervals [] is approximately half a second; and that the speed of motion of the walker's body is approximately constant. There is no difficulty in adding these state propositions, except that the last requires *moving* to have an additional argument, namely the speed of motion. Note that [ti] is the time interval of observation of the walker, and that it contains t, the time at which X is said to be walking. Thus walking is defined by behaviour in the temporal vicinity of the moment of predication, specifically about two seconds of motion allowing about three or four steps.

Our representation of walking is limited since it is not applicable to unusual modes of walking (for example, on hands and knees) or to animals. This limitation raises the question: how many kinds of walking should be represented separately? Also, is there a representation which expresses the common features of all kinds of walking? We have attempted such a representation in Figure 5-5. The representation is based on the following characteristics of walking in general:

1. *it is done using limbs that are a subset of the limbs of the individual involved in the walking;*
2. *the number of limbs involved is greater than or equal to two;*
3. *at all times some of the limbs used for walking are in nonsliding contact with the walking surface (not the same as saying some of the limbs are in contact with the surface at all times);*
4. *each limb used for walking is stationary on the walking surface at some time and subsequently is moving for some time; and*
5. *the individual as a whole is in motion in the walking direction*

The interesting feature of our representation is the use of quantification to describe the role of any number of legs in the walking. Note that without quantification, describing the locomotion of say, a millipede would be very tiresome.

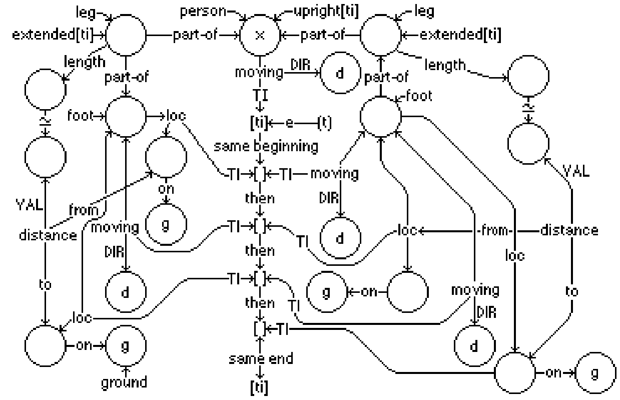


Figure 5-4. Person x walking at time t in direction d on ground g.

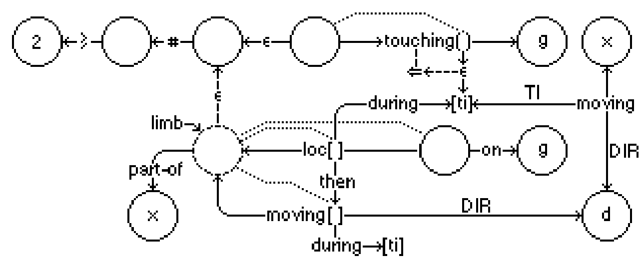


Figure 5-5. x walking at time t in direction d on surface g.

A serious flaw in our representation of walking, and one for which we have no systematic remedy, is that we have ignored the fuzziness of many of the meaning components. For example, it seems necessary to put constraints on the length of stride (lest the walker be allowed to mince forward in millimeter increments), yet to give an exact distance would be absurd.

An important consequence of conceptual fuzziness, considering complex concepts, is that we can no longer draw a sharp boundary between extracting the meaning of an utterance and making probable inferences on the basis of the derived meaning structure. This is because we only find the probable meaning of an utterance. For example, the utterance “John built the house” probably means that he built a large, rigid-walled enclosure with a roof, separate rooms, etc.; but none of this is certain. The utterance “John was laughing” probably means that he was producing a series of voiced sounds by staggered exhalation of air, and that his facial expression was merry; but he might have laughed silently, or his facial expression might have been derisive or even hostile. If we try to reduce semantic uncertainty by excluding from the meaning of a term all but its absolutely minimal content, and ascribe everything else to inference, we run into problems. In the case of house all that would remain would be a partial enclosure - which accommodates a fenced-off field, a shipping crate, or a jacket. In the case of laughing we would perhaps be left with spasmodic breathing and intent to convey amusement, which could suggest that John is asthmatic and dancing a jig.

Finally we wish to point out that many concepts can be understood in different ways. For example, in “John was listening to the incessant chirping of the crickets,” is chirping understood simply by its correspondence to a particular auditory sensation, or is it understood as a rapidly fluctuating, more or less uniformly high-pitched sound, or even as a complex variation of air pressure with time? Minsky's (1975) work on frame systems strongly suggests that the kind of understanding of a concept we use at a given time is extremely task -and-context dependent. This certainly casts doubt on the one-concept-one-formula approach to language understanding.

5.4 Adjectives and Relative Terms

Adjectives and relative terms are typically represented by drawing on a reference set. The notion of a reference set (a set of objects whose members are used for comparison with some given object relative to some measurable attribute of the objects) is difficult to comprehend. While it is possible to define a more or less adequate reference set to account for “a large apple,” it is not immediately apparent what the reference set would be if one were to ask a child to draw a large circle on a sheet of paper. This is a question above and beyond the one pointed out (rightly) by Bartsch and Vennemann (1972) concerning how the reference set is inferred from the context, especially extra-sentential context.

We avoid some difficulties of having predetermined reference sets by making use of functors.[1] The typical value functor applied to a concept with some measure attribute returns a value, for example, the typical value of size for man. Note that this is not the same as the typical man's size. The typical man's size is not readily determinable since it is hard to ascertain exactly what constitutes a typical man. A typical value functor was shown earlier in Figure 3-6. We can abbreviate the typical value functor in a manner analogous to the collapsing of predicates in the abbreviated network notation.

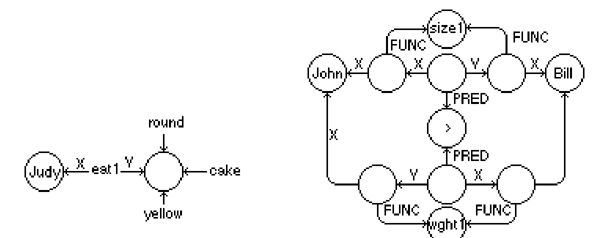
Descriptive adjectives are treated as conjoined predications in most cases, as shown in Figure 5-6a. Yet most adjectives appear to be comparative in nature regardless of their morphology. For example, big, small, tall, heavy, and so on are relative adjectives based on some measurable attribute of the object of focus. Figure 5-6b illustrates how comparatives would be diagrammed. The explanatory paraphrase of “John is bigger than Bill” is “John's size is greater than Bill's size”. Often the comparative is implicit in the utterance. For example, in the sentence “John is a big man” the adjective “big” serves as a comparative. The associated paraphrase is “John is a man and the size of John is greater than a typical value of size for a man”.

Ordinary discourse admits constructions such as: “John is the perfect man.”, “Mary is the worst conceivable cook.”, “In order to form a more perfect union...”, ... Modifiers such as perfect, ideal, and worst conceivable are problematic to represent because of the way they operate on what they modify. For example, we might formulate “John is the perfect man” in logical terms as:

$$(\forall P)[[(\forall x)[man(x)\ \&\ P(x)*?\ y\text{-}approves[P(x)]]]\ ?\ P(John)$$

where *? stands for “necessarily implies”.

where y is the speaker. The formulation reads “John has all properties such that y would approve of any man's having them”. We can then easily formulate an expression for “someone is not a perfect man” by utilizing our formulation given above with the existential quantifier added $(\exists z)\sim$ and replacing P(John) with P(z). Clearly, the method of handling comparative adjectives such as big, tall, etc. does not work here.



(a) Judy is eating the round, yellow cake. (b) John is bigger than Bill but Bill is heavier than John.

Figure 5-6. (a) Mixed bag of adjectives network. (b) Explicit comparative network.

We make no definite proposals for handling adjectives such as perfect, ideal, worst kind of, best conceivable, etc., at any detailed level of analysis. A more superficial analysis of “Big John is a perfect fat man” is rendered as Figure 5-7. Additional information such as “John is a basketball player”, can be easily added to the structure.

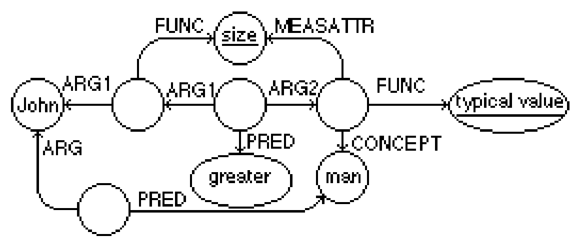


Figure 5-7. Functor Networks: Big John is a perfect fat man.

5.5 Adverbial Constructions

We draw attention to the major problem of representing adverbial meanings and suggest plausible methods for handling adverbial constructions within the state-based conceptual framework.

Two major approaches treating adverbial modifiers include that of Montague (1972, 1974) and Bartsch & Vennemann (1972) who regard comparative adjectives and adverbs as operators which transform predicates, and that approach due to Reichenbach (1947), seemingly accepted by Schank[2], regards adverbial modifiers as second-order predicates that impose constraints on a specific relation, thereby restricting the class of specific relations to which it may be a member. We consider Bartsch and Vennemann's approach, which seems promising but will be seen to have serious de-

---

[1] According to Cresswell, a functor is a symbol which, occurring as the first member of a sequence of symbols of certain syntactic kinds, makes a sequence of the same or another syntactic kind, Cresswell (1974).

fects.

Bartsch & Vennemann suggest that adverbial adjectival modifiers operate on noun meanings; they have semantic representations with functors f such that f is applied to term x to map x onto a new term f(x). The contrast between "Ed owns a large car" & "Ed is running quickly" illustrated a problem with this approach. Whereas large in the first sentence has as a reference set the set of cars, and Ed's car is large relative to the average for that set, running quickly cannot be analyzed so easily. If the analogy were perfect then the reference set operated on by quickly would be the set of runnings (whatever that means); but clearly this set of runnings must be further restricted to the set of runnings Ed is capable of performing. Thus quickly appears to operate not on running alone, but on Ed running.[1]

Thus the nature of the runner is being used to narrow the reference set to which we apply a measure function. In Ed is running quickly - quickly modifies running with respect to Ed's runnings, or, if we don't know Ed, at least to human runnings (assuming that Ed is human). Unfortunately factors other than the identity or category of the runner can also affect the meaning of quickly, as shown by "Ed is running quickly on his hands and knees", "Ed is running quickly on the moon", "Ed is running quickly in Chile", "The cheetah is running quickly in the dense forest", "The cheetah is running quickly on the plain" The effect of locale on the meaning of quickly is seen in the contrast between the first two of these examples and between the last two of these examples.

The context which determines the meaning of an adverbial modifier cannot be circumscribed once and for all. In general, adverbials must be allowed to interact with any specific and general knowledge available about the participants in (and setting of) an action. In the approach of Zadeh (1972) to the treatment of adverbial hedges he specifies (weighted) components of each fuzzy term on which a hedge may operate once and for all. Because he needs to specify these (weighted) components prior to using a particular hedge, his approach lacks generality. In our semantic network, we would represent "The sun is shining brightly on the beach" without the adverb as diagrammed in Figure 3-6, in keeping with Bartsch & Vennemann's general approach but taking into account the above considerations. In the representation we show the explicit relationship between the speed of the sun's shining as compared to the typical value of brightness for something that is shining on a beach and is a sun.

It is well to note that the set of suns shining brightly on beaches required for comparison, may well be empty (if not, replace "beach" with "ocean floor"). The reference set therefore, if it exists at all, is not of this world but of some imaginary world which is our conception of how hard suns would find the going if they were to shine brightly on beaches (or ocean floors). In our formulation we have applied the typical value functor to the lambda abstracted predicate "shining brightly on the beach". The typical value functor does not presume the existence of a reference set.

5.6 Opacity and Vagaries of Reference

Some linguistic forms give rise to referentially opaque contexts. This is true of the propositional attitudes "believes that ...", "knows that ...", "wants to ...", and others, as well as other modalities created by causal situations, intentions, and the like, and conditional statements including the counterfactual conditional. Quotation creates referentially opaque contexts. While "simpleton" may be referentially equivalent to "fool", the statement simpleton has nine letters does not allow substitution of "fool" for "simpleton".

To illustrate how a referentially opaque context can block existential quantification, the sentence "John wants to marry a blonde." gives rise to two possible interpretations:

"John wants to marry a specific girl who also happens to be a blonde."; and

"John has no particular girl in mind, but he wants whoever he does marry to be a blonde."

The first interpretation, transparent reading, can be existentially quantified, that is, there exists someone whom John wants to marry. The second interpretation cannot be quantified in like manner since it contains an assertion about an existential statement rather than being an existential statement.

Various (equivalent) explanations have been given for this type of ambiguity. Philosophers tend to describe this as scope ambiguity of an existential quantifier. Some linguists however, prefer to portray the ambiguity as a distinction between a referential and attributive use of a noun phrase. With Schubert's notation the opaque reading of "John wants to marry a blonde" would be represented as shown in Figure 5-8. The transparent reading would be represented by Figure 5-8 if we took out the dotted line.

Both Montague (1972) and Lewis (1972) have developed theories that enable both the transparent and opaque readings for sentences to be generated. This is not carried far enough. The important problem remaining is how to choose the correct interpretation in context. This problem is investigated further in Strzalkowski & Cercone (1986, 1989) and a solution is proposed to choose the correct context.

---

[2] Schank diagrams adverbs as action modifiers without further analysis. Apparently he has not concerned himself with the meanings of genuine manner adverbials so far, however, see Schank (1974) for a discussion of adverbs such as vengefully, thoughtlessly, etc. In the case of many adverbs (as in the case of many adjectives) this neglect is probably justified, since most of the meaning content derives from perceptual processes. For example, in the sentence Mary walked gracefully it is difficult to paraphrase gracefully in more elementary terms. Essentially we know gracefulness when we see it. Perceptual understanding needs to be supplemented only by a few additional facts for language comprehension, such as the fact that graceful motion is generally pleasing, is more or less the opposite of awkward motion, is smooth and coordinated, etc. Other adverbial modifiers clearly require systematic analysis; quickly is a good example. This term appears to say something about the speed of an action or activity, comparing it to some standard. An adequate meaning representation for quickly should spell this out precisely.

[1] cf "The cheetah is running quickly" and "The ant is running quickly" Clearly quickly here operates on running ant and running cheetah respectively.

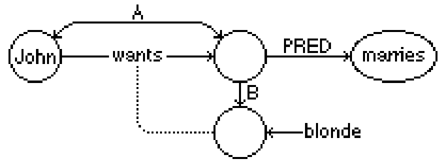


Figure 5-8. John has no particular girl in mind, but he wants whoever he does marry to be a blonde.

We have taken a critical look at many problems and suggested plausible methods for dealing with some of them. In particular, the basic representation in terms of states and events, the definition of complex concepts (most importantly action concepts), the handling of adjectives and relative terms as well as adverbial modifiers were examined. Solutions to the problems that these topics present to language processing systems have been developed to the point that some of them have been incorporated into the experimental programs that support this research.

6.0 Superimposing Organizational Strategies

6.1 Representing Lexical Info: English Word Meanings

In Quillian's networks word meanings were represented by a network of objects and relations among the objects. To distinguish ambiguous meanings for pairs of words Quillian compared the plausibilities of different interpretations by the strength of the associations linking the various pairs of meanings. Quillian's network was believed to contain the germina for a more sophisticated understanding of the relationship between meaning and grammar. Subsequent writers who used nets in their systems tried to further explore this relationship.

Schank's conceptual dependency theory represents meaning structures with a graphical notation consisting of items from four conceptual categories: picture producers [PP], picture aiders [PA], action aiders [AA], and actions [ACTS]. They correspond closely to nouns, adjectives, adverbs, and verbs, respectively. The conceptualization is Schank's smallest structural unit; conceptualizations are graphical structures that link together conceptual categories using a variety of graphical symbols, conceptual tense markers, conceptual cases and primitive actions.[1]

Cercone augmented the meaning representations of Schank and Wilks, utilizing semantic networks for the meaning representation of both the semantic and pragmatic information of a word concept. Cercone's meaning representation is not based on primitives but permits efficient use of semantic preferences and is capable of accommodating unlimited amounts of information about complex concepts without loss of computational efficiency in the use of those concepts. The pragmatic and semantic information associated with the concept drink1, the ordinary sense of drinking as in John drinks water, is illustrated in Figure 6-1. Note the constraints that the drinker should be animate and the stuff drunk liquid. The major implication that x ingests y (or the subsuming concept ingest) in turn provides access to the implications of ingesting. In this way Schank-type and Wilks-type inferences are made available through property inheritance. Figure 6-2a shows how the ordinary sense of drinking can be modified to accommodate supplementary propositions to explain the implications associated with an alcoholic drinking.

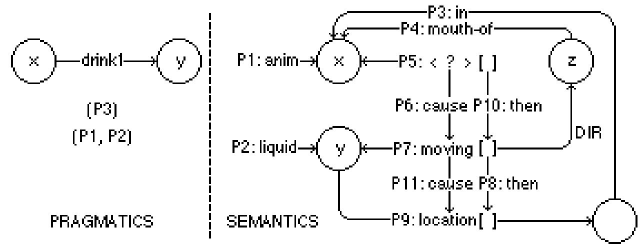


Figure 6-1. The properties of drink1.

Since ingest is not a primitive, it also has associated pragmatic and semantic properties. These are illustrated in Figure 6-2b. A paraphrase of the semantic formula says that if x ingests y through z, then some unspecified state or event causes stuff y to move towards the opening z and this in turn causes y to assume a location inside x (without trying to be very sophisticated about this point).

6.2 Organizing Network Information

The characteristic concept-centred organization of semantic networks does not address representation issues but rather focuses primary concern with organizing knowledge for effective use. Subsequent semantic network notations have been developed in an independent and application specific manner. Moreover, they have often blurred the important distinction between the representational and organizational aspects of network formalisms.

Early efforts by Shapiro (1971) to imbue networks with increased logical power explicitly documented this distinction by contrasting system relations, items, and item relations. Schubert clarified this distinction by demonstrating that a logical representation couched in network form offers the advantages of a propositional representation (that is, formal interpretability and expressive adequacy) while retaining the methodological advantages of the associative network organization. His notation indicated that an intelligent indexing scheme coupled with a database of logical formulae could indeed be considered to be a kind of semantic network.

The basic distinction between the propositional content of a knowledge database and the access mechanism to that content has been noted by Bobrow and Winograd (1975). We wish to emphasize organizational aspects of semantic net-

[1] Schank used four cases: the objective case, which relates an objective PP to an ACT; the recipient case, which relates a donor PP and a recipient PP to an ACT; the directive case, which relates direction (to and from) to an ACT; and the instrumental case, which links conceptualizations instrumental to an ACT to a conceptualization containing the ACT. In addition to conceptual cases, Schank makes use of only fourteen primitive actions through which he expresses all other actions. These primitive actions are: PROPEL, MOVE, INGEST, EXPEL, GRASP, PTRANS, MTRANS, ATRANS, SMELL, LOOK-AT, LISTEN-TO, CONC, and MBUILD.

works, in the tradition of Quillian and in the spirit of Hayes (1974), who writes, "If someone argues for the superiority of semantic networks over logic, he must be referring to some other property of the former than their meaning". The correspondence between semantic networks and logic has been established; the meaning of a given network is identical with the meaning of the equivalent logical expression. The object of our immediate attention is that structure which remains after paring the propositional content from a semantic network, that is, the indexing structure which provides concept-based proposition access.

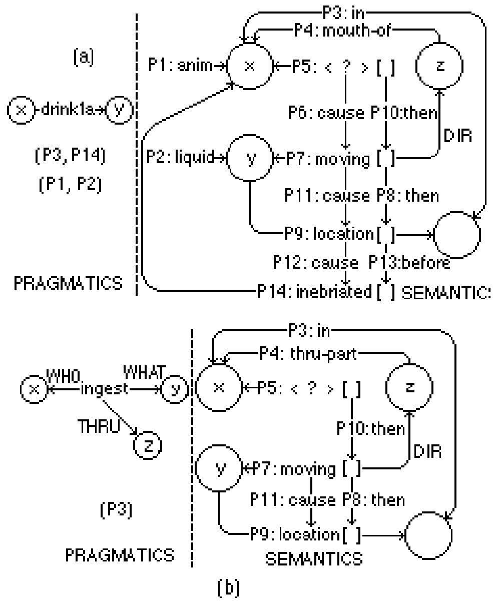


Figure 6-2. (a) Properties of drink1a. (b) Properties of ingest.

Organizational theories of knowledge can be characterized by the desire to cluster related knowledge into chunks. Ideally, these chunks should reduce the computation required to isolate knowledge relevant in a particular context.

We can easily impose a hierarchical (subconcept-superconcept relation) structure on top of the general concepts in memory, such as that illustrated in Figure 6-3, as a heuristic device. Earlier on, both the concept-centred organization of nets and many of the logical tools of predicate calculus were evident (albeit implicitly) in many of knowledge representation systems. For example, in reference to the GUS system (Bobrow et al., 1977), Kay (1976) reports " ... now the contents of these slots in the dialog frames (and in lots of other frames that exist in the system) are typically other frames. These structure recurse to great depth. Of course they are not simply tree structures, but they are circular and point to one another; they're networks". Also, Hayes provides a translation of KRL features into a many-sorted predicate logic, which he takes to be the external meaning of KRL expressions.

The conspicuous remaining feature of frame-like systems is simply the idea of grouping pieces of knowledge which may be useful for understanding a particular concept or situation. Hayes explains that a frame may be viewed as an n-ary relation between itself and its slots, which themselves may be viewed as binary relations and unary predicates. One could therefore represent a frame within the semantic network notation. The major difference between the "frames" view and the network view is one of function versus structure, as noted in Schubert et al. (1979): "A memory structure is regarded as a frame because of the kinds of knowledge and capabilities attributed to it, rather than because of any specific structural properties".

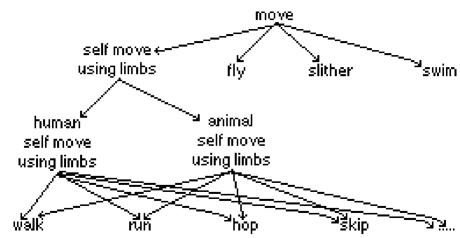


Figure 6-3. Superimposed Hierarchical Structure.

6.3 Superimposing Topical Organizations

Research into semantic networks at the University of Alberta culminated in one solution to the so-called symbol mapping problem and in doing so, this group of researchers directly addressed organizational aspects of semantic network representations, see Schubert et al. (1979).

In general a concept is characterized, though not defined, by its implications and these implications can be more or less essential to the meaning of a concept. For example, consider the following two statements: Johnny walked his pet boa-constrictor daily and Johnny's boa constrictor has six legs or Boa constrictors are friendly, warm furry animals. Both statements contradict boa constrictor properties, the latter alternatives contradicting an essential property. In the second statement, we are either violating necessary universal statements for parts (has six legs) or properties (friendly, warm, furry); in the first we are merely contradicting contingent properties (pets).

It is clear that any system designed for reasoning about its world must efficiently exploit property inheritance within generalization (IS-A) hierarchies or relationship inheritance from components to corresponding components (PART-OF) hierarchies. What complicates this problem is that conceptual entities typically consist of many components, for example, parts of an object, the participants of an action, or the departments of an organization. An example illustrates one possible solution. Consider a bird subhierarchy; it should be sufficient to specify the attachment relationships between head, neck, body, legs, and tail at the top level, and this information should be visible from each particular kind or instance of bird. Consider the following fragments of bird knowledge:

(∀r) [[r bird] ? (∃s) (∃t) [[s part-of r] & [t part-of r] & [s head] & [t neck] & [s joins t]]]

(∀u) [[u owl] ? (∀v) (∀w) [[v part-of u] & [w part-of u] & [v head1] & [w neck1] ? [v big] & [w short]]]

(∀x) [[x emu] ? (∀y) (∀z) [[y part-of x] & [z part-of x] & [y head2] & [z neck2] ? [y small] & [z long]]]

We assume that in addition to the subconcept relationships that owl, emu are necessarily birds, that head1, head2 are necessarily heads, etc. Particularized owl's heads, etc. are partly intuitive (the picture it conjures in the mind) and partly anticipatory. We will need a separate concept for each part of a thing as a point of attachment for knowledge peculiar to it. This disjointed collection of propositions is redundant. Furthermore, nontrivial inference is required to transfer relationships from the bird context to the owl and emu contexts. Here the only relationships, besides part-of relationships, is that the head is joined to the neck. However, there would be many such relationships in a system knowledgeable about birds in general.

Consider the alternative arrangement of these facts:

(∀x) (∀y) (∀z)

[[[x bird] ? [[y part-of x] & [z part-of x] & [y head] & [z neck] & [y joins z]]]

& [[x owl] ? [[y head1] & [z neck1] ? [y big] & [z short]]]

& [[x emu] ?[[y head2] & [z neck2] ? [y small] & [z long]]]]

Through variable sharing we have eliminated all redundancies. Moreover, parts relationships for birds now transfer trivially to owls and emus. Thus shared *nodes* can change character depending on viewpoint. In our bird propositions x represents any bird from one view, an owl from another, and an emu from another. Similarly, y and z represent different heads and necks depending on the point of view.

Back-linking from shared variables to propositions should be suppressed, since there is no benefit in having uniform access to all propositions in which such nodes participate. It is more useful for general knowledge to be accessible via participating predicates, such as “owl”, “head”, etc.

We can generalise from our example and conclude that the knowledge associated with a generalization hierarchy should be stored as a set of implicative propositions sharing one universally quantified node and any number of existentially quantified nodes dependent on the universally quantified node. The antecedents of the implications involve the universally quantified as argument, and correspond to the concepts making up the generalization hierarchy.

Unfortunately, the shared variable form of generalization hierarchies complicates the process of adding new information. New facts such as (∀x)[[x owl] ? [x predator]] cannot be simply *added* to the net by creation of a new variable node. Instead, this information must be inserted at the appropriate place in the appropriate hierarchy, with x replaced by the universal node of that hierarchy.

We have only considered monadic concept hierarchies and it seems possible to organise relational concepts hierarchically as well. These will share more than one universal node, as exemplified with the following fragment of the “ingests” hierarchy [We have suppressed time relations and other subtleties for simplicity, but see Section 7 below]:

(∀x) (∀y) (∀u) (∀v) (∀w)

[[[x ingests y] ? [[u orifice-of x] & [v place] & [w place] & [v outside x] & [w inside x] & {y moving v w] &...]]

& [[x eats y] ?[[y food] & [u mouth-of x]]]

& [[x drinks y] ? [[y liquid] & [u mouth-of x]]]

& [[x snuffs y] ? [[y powder] & [u mouth-of x]]]

:

etc.

How many hierarchies are there? We imagine that the most hierarchies should be of the order of generality of Schank's or Wilks' primitives.[1] Thus there may be dozens of hierarchies.

Symbol mapping in a semantic network is facilitated by imposing a sub-concept super-concept (IS-A) hierarchy on the network concepts. Thus Clyde's elephanthood immediately provides a handle on knowledge attached to the elephant concept. There is more at stake with Fahlman's “Clyde the elephant” than property inheritance. Mere access to elephant knowledge does not guarantee swift question-answering or consistency checking. Imagine hundreds of facts impinging on “Clyde”, “elephant”, “mammal”, etc. and attempt to do a Quillian-like activation search to particular attributes, such as colour or appearance. This would, in all likelihood lead to a combinatorial explosion when trying to construct inference chains to answer relatively simple queries like “What colour is Clyde?” or “Does Clyde live in a teacup?”.

To answer questions of the sort just posed, we note two features which these examples illustrate. One is the need to classify propositions topically as colour propositions, location propositions, size propositions, etc. This classification scheme should help us to avoid the exhaustive search for combinations of propositions which yield a desired conclusion. The other is the need for access to just those propositions about a concept which belong to one of the above topics.

Our approach is to structure the propositions associated with each concept in accordance with a topic hierarchy. We define a topic as a predicate over proposition-concept pairs. For example, “colouring” is a predicate which is considered to be true for the proposition “a zebra has black and white stripes” in relation to the concept “zebra”. Another topic predicate which is true for that proposition in relation to “zebra” is “appearance”, in fact, “appearance” holds for any proposition-concept pair for which “colouring” holds, that is, “appearance” is a supertopic of “colouring”, and conversely, “colouring” is a subtopic of “appearance”.

Topic predicates are stored in the semantic network, linked by subtopic and supertopic relationships. Together, these form a topic hierarchy (or several topic hierarchies). Topic hierarchies provide a basis for organizing the propositions attached to a node for a particular kind. A possible topic hierarchy for physical objects is shown in Figure 6-4, which provides an attempt to comprehensively classify knowledge about physical objects with minimal overlap between categories. The subconcept topic is intended to be a slot not only for genuine subconcept relationships (that is, necessary sub-

[1] This may be the real significance of primitives.

sumption) but also for contingent subsumption relationships and for instances of a concept. Similar notions apply for superconcept relationships.

Once a topic hierarchy has been defined for a particular kind of node, the propositions attached to any node of that kind can be organised in accordance with the hierarchy. This is accomplished by superimposing an access structure called a topic access skeleton upon the attached propositions. A topic access skeleton *mimics* a part of the topic hierarchy, namely that part which is needed to supply access paths to all the available propositions about the node, when these are attached to the appropriate terminal topics.

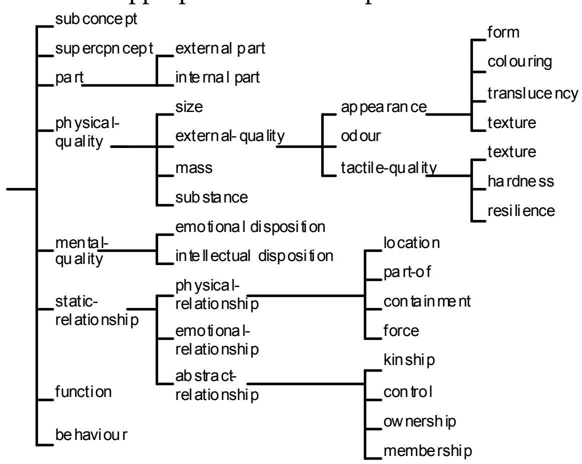


Figure 6-4. A topic hierarchy for physical objects.

For example, if the only facts known about Clyde are that he is an elephant and likes to eat peanuts, these would be attached to the access skeleton. If elephants, in turn, are known to be very large, grey, rough-skinned mammals and Clyde is known to be an instance of an elephant, these facts would be attached to the access skeleton. Note that “texture” appears twice, making the “rough-skinned” predication available both as an aspect of appearance and as a tactile quality. In implementations a topic hierarchy and corresponding access skeletons need not be strictly tree-structured since a single “texture” node can be used, with pointers to it from both the “appearance” and “tactile-quality” nodes.

Schubert et al. (1979) discuss the insertion and retrieval of propositions in a topically organised semantic network in detail. In that discussion they indicate how topically organised networks facilitate the kinds of inferences about objects and their kinds and discuss the importance of the subconcept superconcept classification. They also explain a mechanism for performing automatic topical classification of propositions and discuss how the proper topical classification of propositions in general depends on their logical form and on the nature of the predicative concepts involved. Finally they address time and storage trade-offs and provide a clever path contraction algorithm which guarantees descent time in the tree-like topic hierarchies, and subsequent access to propositions they encode, to be proportional to log nmax, the maximum number of propositions attached to any concept.

7.0 Special Purpose Inference Mechanisms

Topic hierarchies are obviously useful to organize the retrieval of information relevant to the implications of concepts, but the same idea can be extended to other kinds of information about concepts. For example, time, part and colour information about concepts can be organized into structures which support the efficient use of that information. Another way to view the incorporation of this organizational knowledge is simply to view it as a special purpose inference system augmenting an ordinary first order logic (FOL) theorem prover. From this view, the topic hierarchy is simply a specialist about the implications of being a concept instance. This view has given rise to the method of creating hybrid reasoners in which a general purpose FOL reasoner is augmented by special purpose methods that can efficiently reason about special relations in particular domains (Slagle, 1972; Bundy, Byrd and Mellish, 1982; Vilain, 1985; Brachman, Gilbert and Levesque, 1985; Rich, 1985). The ECO system is one of these hybrid systems, supplementing its general purpose reasoner with special purpose methods for time, types, numbers (arithmetic relationships), sets, colours, strings (for string manipulation), and part-of relationships.

The key to the efficiency of many of the specialists is their use of alternate representations, which enable the reasoning to be done by efficient algorithms that take advantage of special properties of the predicates, terms and functions in their domain. For example, the temporal specialist uses a partial order graph to represent temporal relationships and fast graph searching techniques to determine the relations, the type specialist uses a preorder numbering scheme on type hierarchies, and the colour specialist uses a cylindrical model of colour to determine relationships among colours.

A significant problem for designers of hybrid systems is determining how to integrate the special reasoners into the general reasoner. Generally, the methods fall into one of two classes: those which operate at the metalevel, using their special abilities to improve the order in which available inference steps are applied, and those which operate at a *sub*level, seeking to improve the power and efficiency of the inference steps used by the general reasoner. In the ECO system, all of the specialists are added as sub-level reasoners. At the sub-level, there are (at least) four ways in which a specialized reasoner can be added: (1) as an *evaluative* inference step, which directly evaluates truth values of literals, or simplifies functional terms; (2) by changing the unification procedure of the inference step to take into account some of the specialist's knowledge (e.g., to disallow unification of variables that are typed with incompatible types, or to unify two constant symbols that in fact stand for the same domain concept); (3) by expanding the set of literals which lead to the possible application of an inference step (e.g., in a resolution system a colour specialist might immediately identify white(x) and black(x) as incompatible clauses, or identify man(x) as being subsumed by human(x)); and (4) as a completely new inference step (e.g., the paramodulation step that is commonly added to resolution systems to incorporate equality). Specialists in the ECO system can use any of the first three methods to integrate their knowledge back into the general reasoner. The details can be found in Schubert, Papalaskaris and Taugher (1987), which also discusses the relationship of (1) and (3) to Stickel's theory resolution (Stickel, 1985). Aside from the specialized domain knowledge that is encoded into the specialist, specialists generally acquire knowledge by transforming statements in the language of the general reasoner into their own representational forms. Providing a mechanism which allows specialists to communicate with the general reasoner is only part of the solution to increasing the overall efficiency of a hybrid reasoner; as Levesque and Brachman (1987, p 88) put it, "the trick with these hybrid systems is to factor the reasoning task so that specialists are able to cooperate and apply their optimized algorithms without interfering with each other." Some steps towards solving this particular problem within the ECO system framework are described in (Miller and Schubert, 1988b). In contrast to hybrid systems in which the specialists add expressive or inferential power to the overall system, for example, KL-TWO (Vilain, 1985) and KRYPTON (Brachman, Gilbert and Levesque, 1985). ECO specialists employ alternate representations of the same knowledge available to the general reasoner, and serve only to accelerate inference. Thus, there is no requirement that a specialist be complete, as the general method can fill in any gaps, albeit less efficiently. As long as the operations a specialist is allowed to perform are equivalent to sets of standard deductive steps, the specialist (and thus the overall system) is guaranteed to be logically sound.[1] The following sections describe the representational and inferential capabilities of several of the more complex specialists used by the ECO system.

7.1 Time

The time specialist described here is based on the specialist designed by Taugher and Schubert (Schubert, Papalaskaris & Taugher, 1987), and includes some enhancements to handle both strict and nonstrict relations, and combinations of events, time points and absolute times in propositions (Miller and Schubert, 1988a). The representation used is a partial order graph that has been partitioned into *chains*. All the points belonging to a chain are linearly ordered with respect to each other. There may be transitive arcs between the points in a chain. Cross chain links define relations between points in one chain and points in another.

For points within a chain, an arbitrary pseudo-time number is associated with each point (a minimum and maximum). These numbers show the ordering relationship between points in a chains. In addition, the minimum point and the maximum point on the chain that a point can be equal to are stored with it-giving a range of points that can possibly be equal. These are used to show whether the relationship given by the pseudo times alone is strict or nonstrict (for example, < or < =). Determining the relationship between any two points in the same chain can be done in constant time using these pseudo-times, while a graph search is required if they are on different chains.

In addition to the time graph of time points, there is a metagraph of chains. The cross chain links define arcs between chains in the metagraph. The metagraph is used to search for paths from one point to another. This makes a graph search dependent on the number of cross chain links rather than the total number of time points (a significant savings).

Figure 7-1 shows an example time graph and meta graph. In the time graph, small circles represent points on chain 1, small squares are points on chain2, and small triangles represent points on chain 3. So circle1 is before circle2, circle2 before circle3 and so on. There are cross chain links from triangle1 to circle1 (that is, triangle1 is before circle1), from triangle1 to square1, from square2 to circle1, and one more from circle2 to square3.

In the metagraph, these cross chain links show up as links between meta-nodes. There is one meta node for chain 1 (the big circle), one for chain 2 (the big square) and another for chain 3 (the big triangle). The links within chains do not show up here, as within a chain they are not needed to determine relations. Following the cross chain links, we can get that triangle1 is before square3, and square1 is before circle3, but no information about triangle2 and square3.

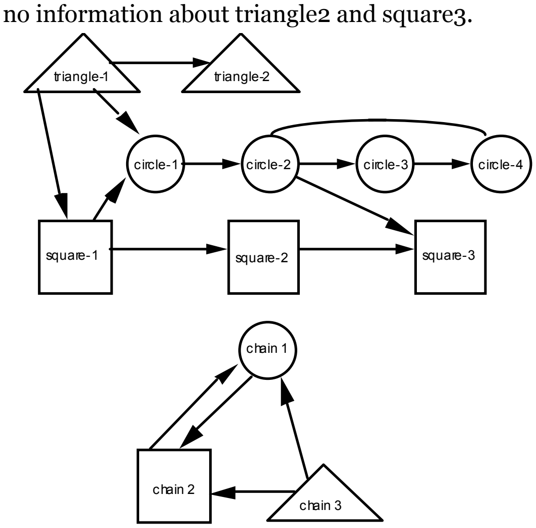


Figure 7-1. Example Timegraph and Metagraph.

Furthermore, an absolute time (date) minimum and maximum are stored with each time point. These are a six-tuples of the form <year, month, day, hour, minute, second>, where each element may be numeric or symbolic, for example, 1987 04 a 12 b c represents some time at or after 12 a.m. and before 1 p.m. of some day in April, 1987. Symbolic information may be filled in later by another assertion, or left unspecified throughout the session. Absolute time maxima propagate back to points before the given point (in the chain or on other chains), and minima propagate forward. This ensures that each point has the best absolute time information possible. Absolute time comparisons can sometimes be used to get a relation in constant time between two points on dif-

[1] This restriction is satisfied by the operations ECO specialist are allowed to participate in.

ferent chains, avoiding a metagraph search.

Insertion time into the graph is constant in most cases, except for propagation of absolute times or strictness values. In the worst case, propagation may require visiting every point in the graph, although this is highly unlikely. Occasionally a chain may have to be renumbered, which requires going to all the points in a single chain.

Creation of all supporting graph structures requires O(n+e) space and O(n+e) time, where n is the number of time points, and e is the number of relations between them. Determination of relations between points is based the metagraph, as the in-chain checking time is constant, so is O(m) where m is the number of cross-chain links (m < < n sup 2). Duration minima and maxima (in seconds) are stored on the links between points. These may affect the absolute times around them, which are then propagated. They are also used in calculating duration between points where the path uses this link. Durations may be unspecified and are then treated similar to unspecified absolute times, generating an evaluation request and adding to the interested party list of the concept. To determine the duration between any two points, an exhaustive search must be done between those points, calculating the duration along all paths to get the best one. This particular search uses a traditional depth first search over the entire time graph, rather than using the metagraph. Both duration information on arcs, and duration information implicit in absolute times are used.

7.2 Types

This specialist uses partitioning hierarchies of type predicates as a logically true representation of the relationships among the predicates that appear in the hierarchy. A preorder numbering scheme makes possible constant time determinations of the relationships between predicates within the same hierarchy. For example, in the "thing" hierarchy shown in Figure 6-2, "thing" is divided into "physical-object" and "abstract-object", which are further subdivided: "physical-object" into "living-thing" and "non-living-thing" and so on. The numbers following the predicates are assigned when the hierarchy is created or changed, and can be used to quickly determine subsumption or disjointness relationships. For example, "wolf" and "human" are disjoint because there is no overlap between the numbering range associated with "wolf [38,38]" and "human [17,26]". "Wolf" is subsumed by "creature" because the numbering range associated with "wolf [38,38]" is within the numbering range associated with "creature [16,40]".

The "thing" hierarchy is an "exclusion" hierarchy - all sibling nodes are mutually exclusive. "Overlap" hierarchies are also possible - in these, subsumption can be determined by the same method, but not disjointness. Hierarchies may also be inextricably intertwined, in that the same predicate can appear in several hierarchies, all "connected" at that predicate. For example, if "human" is partitioned in two different ways (minor, adult, etc., as in the "thing" hierarchy; and Caucasian, Asian, Negro, etc., in another hierarchy), it is still possible to determine that Caucasian is subsumed by creature, simply by noting that "human" is subsumed by "creature" in the first hierarchy, and "Caucasian" is subsumed by "human" in the other. This works much like the metagraph in the time specialist - within each hierarchy the preorder numbering is used, and between hierarchies the connecting predicates are used. Thus, we can not only determine relationships between predicates in the same hierarchy, but also relations across hierarchies.

```
THING [1,112]
    PHYSICAL-OBJECT [2,108]
        LIVING-THING [3,40]
            PLANT [4,15]
              ...
            CREATURE [16,40]
                HUMAN [17,26]
                       ADULT [18,20]
                            MAN [19,19]
                            WOMAN [20,20]
                       MINOR [21,26]
                         ...
                ANIMAL [27,40]
                       MICROBE [28,28]
                         ...
                       LARGER-ANIMAL [32,40]
                            FISHLIKE-ANIMAL [33,33]
                         ...
                            WARM-BLOODED-QUADRAPED [37,39]
                                 WOLF [38,38]
                                 FOX [39,39]
                            SIMIAN [40,40]
        NON-LIVING-THING [41,108]
            INANIMATE-NATURAL-OBJECT [42,55]
              ...
            ARTIFACT [56,107]
              ...
            FOOD [108,108]
    ABSTRACT-OBJECT [109,112]
        THOUGHT [110,110]
        IDEA [111,111]
        GROUP [112,112]
```

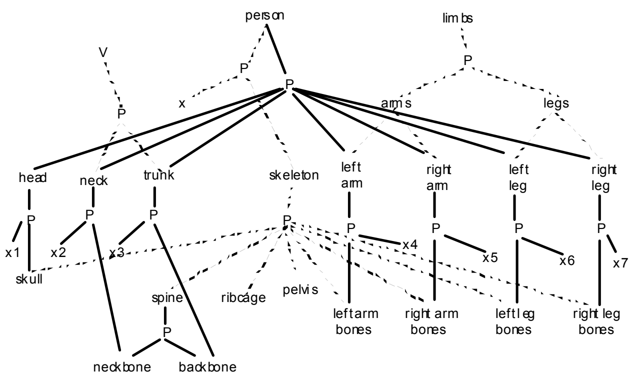
Figure 7-2. A partitioning type hierarchy for "thing".

In general, relationships among type predicates may define an arbitrarily complex graph that cannot be represented by a simple partitioning hierarchy (or even multiple intertwined hierarchies). Nevertheless, the type specialist uses hierarchies because in practice most taxonomies appear to fit reasonably well within a hierarchical framework, see Schubert et al. (1987). More to the point perhaps, is that efficient methods for unrestricted graphs do not appear to be within reach, Schubert (1979). In the ECO system, specialists are not required to be complete, which means that they can trade-off completeness for efficiency without sacrificing the completeness of the overall system.

7.3 Parts

The part-of structure of an object can be represented in essentially the same way as the taxonomy of concept types. For example, Figure 6-3 depicts a partitioning graph which (exhaustively) partitions the human anatomy. The algorithms sketched for type hierarchies could be used to determine the truth values of such formulas as (backbone-of-John part-of spine-of-John) or the incompatibility of (x pelvis-of John) and (x left-leg-of John). However, consider the question "Is the spine part of y?", where "y" is the combination of trunk and neck, as specified in the graph. Since "spine" and "y" are on different partitioning hierarchies, which are furthermore not connected by a part which can supply a transitive relationship between the two (that is, as "human" supplied a relationship for "Caucasian" and "creature"), the type specialist would answer "unknown". However, it should be possible to infer “yes”, since the spine is divided fully into the neckbone and backbone, both of which have upward paths to “y”.

This additional complexity of parts graphs in comparison with type graphs has led to less restrictive structures for representing part-of relationships, at the cost of slightly less efficient algorithms. One such representational structure is a closed graph. A closed graph consists of one main hierarchy, along with any number of other hierarchies such that all of the leaves of these hierarchies are also leaves of the main hierarchy. Figure 6-3 is almost a closed graph; it can easily be made into one by partitioning “x” (which intuitively represents the soft tissue of the body) into “x1” (which represents the soft tissue of the head), “x2”, and so on..



Each P-token represents a partitioning assertion dividing the node to which it is linked above into the nodes to which it is linked below. The solid lines define a partitioning hierarchy, and the broken lines define three additional, superimposed hierarchies.

Figure 7-3. Upper levels of a partitioning graph - the human body.

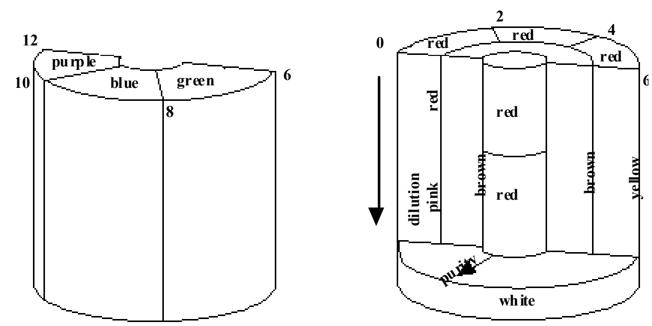
The inference algorithm for closed partitioning graphs (P-graphs) works by “projecting” nodes which do not lie on the main hierarchy into the main hierarchy. For example, the projection of “spine” into the main hierarchy is the set of nodes S = {neckbone, backbone} and for “y” it is Y = {neck,trunk}. From these, it is easy to infer (by using a numbering scheme such as used for type hierarchies) that “spine” is part of “y”, since all members of S have ancestors in Y. If the graph can be decomposed into hierarchies such that no node belongs to more than one of a few hierarchies, and nodes being compared usually belong to a common hierarchy, then expected time for a given comparison will be nearly constant (due to the numbering scheme) since a projection onto the main hierarchy will only be required in a few instances.

The restrictions on the structure of closed P-graphs can be relaxed somewhat, by using semi-closed P-graphs. A semi-closed P-graph is one which is either a closed P-graph, or a semi-closed P-graph with another semi-closed P-graph attached to it by one of its main roots. Intuitively, a semi-closed P-graph allows for “entirely unrelated” partitionings of the same entity. More detail about the parts specialist can be found in Schubert et al. (1987).

7.4 Colours

The colour specialist determines the relationships between colour predicates (for example, crimson is subsumed by red, and brown and blue are incompatible). At first glance, it would seem that the graphical structures used for types or parts would also be amenable to representing colour relationships. However, there are several complications that are peculiar to colour predicates. For example, many colour shades overlap (for example, tan, midbrown, chocolate), and multiple partitionings would be needed to properly separate all of these shades into overlap and non-overlap relationships. Additional partitionings would also be needed to properly represent shades which straddle disjoint basic colours (for example, as lime straddles yellow and green). To deal with “hedged” colour relations, such as “lime is sort of yellow and sort of green”, the partitioning graphs would have to be augmented with adjacency and/or apart-from relations, and even these partitionings would still not be able to represent colour properties such as lightness, purity, saturation, and the warm/cool distinction. As an alternative to graphical representations, it appears that geometric representations are much better at representing these kind of colour relations, and that is why the colour specialist uses the cylindrical colour model depicted in Figure 6-4.

This model, developed by Papalaskaris, in Schubert et al (1987), was arrived at by imagining that any colour is composed of some amount of a pure, monochromatic colour, plus certain amounts of black and white.



The eleven basic colours in a (hue, purity, dilution) colour space (with the cool shades "lifted away").
Purity decreases as black is added to a pure colour, and dilution increases as white is added.
Purity = pure colour/(pure colour + black)
Dilution = white/(pure colour + black + white)
The numerical values have been chosen on purely intuitive gournds. They could be quite drastically altered without affecting the results of the algorithms based on the model, as long as the region adjacency relationships are not changed.

Figure 7-4. Colour cylinder with the cool shades lifted away.

There are three dimensions to this object: (1) hue - this dimension runs through the continuum of rainbow hues, arranged in a circle and arbitrarily scaled from 0 to 12; (2) purity - the radial axis, parametrizes the amount of black present purity = pure colour / (pure colour + black) which decreases from 1 to 0 as black is added; and (3) dilution - axial dimension, parametrizes the amount of white present dilution = white/(pure colour + black + white) which increases from 0 to 1 as white is added.

The model renders each English colour term simply as a region bounded by six coordinate surfaces, defined by three pairs of upper and lower bounds on hue, purity and dilution (so in the implementation each colour is represented by six numbers). With this geometry, it is possible to check any desired relationship between pairs of colour regions, such as inclusion, overlap, adjacency, and separation by using the corresponding geometric concepts. Hedged predicates are handled by simply expanding the colour regions of the colours in question.

8.0 Current Research and Directions

Research and development within the ECO family endures; we are trying to achieve the goal of developing an English language conversational system. Recent work on the inference engine has shifted our focus from a resolution-based predicate logic question-answering system, see deHaan and Schubert (1986), to an inference engine which uses natural-deduction-like inference steps and is based on a new episodic logic which allows both the explicit content of narratives or dialogs and the world knowledge needed to understand them to be easily represented. This logical representation provides restricted quantifiers, modal operators and propositional attitudes, predicate modifiers, nominalization operators, episodic variables, anaphoric variables, unreliable generalizations, and other non-standard constructs, see Schubert and Hwang (1988) and Schubert and Hwang (1989). The new inference engine, named EPILOG, also performs input-driven inferencing to generate "interesting" conclusions before they are actually required to answer a question. Specialists may participate in this process, by making assertions back into the knowledge base of the general reasoner. Meaning postulate "axiom schemas" are supported, for example, a meaning postulate might assert that any concept that is described by a predicate modified by "very" can also be described by that predicate standing alone. Meaning postulate axiom schemas greatly reduce the number of rules required. A specialist to handle the "meta-predicates" that appear in these meaning postulates (for example, "action-pred") has been added, as have specialists for strings, arithmetic relationships, and set relationships, see Schubert et al. (1990). Work continues on the conversational system and on EPILOG at the Universities of Rochester and Alberta, and also at Boeing Co., where EPILOG is an important part of a prototype message processing system, see Jenkins et al. (1990).

9.0 Concluding Remarks

A little over ten years ago, we made a number of remarks about future directions, see Schubert et al. (1978, 1979). It is instructive to briefly review some of those remarks now. We said, Schubert et al. (1978, p 170),

> "... an important future task will be the integration of uniform propositional representations with special-purpose representations, such as those required for efficient spacial, linguistic, and numerical information processing."

I think it is safe to say that this integration has largely taken place, see deHaan and Schubert (1986), Miller and Schubert (1988a, 1988b), & Schubert et al. (1987). We went on to say, Schubert et al. (1978, p 170),

> "... we need to transplant the parser to the topically organized net and expand it to handle at least noun phrase reference and bring into play the 'major implications' of verb concepts."

Much work has been done on parsing since that statement was made, both in Alberta and by derivative groups of researchers, see Strzalkowski (1983), Strzalkowski and Cercone (1986, 1989). Although, strictly speaking, the words of our quote have been performed, the intent may well remain elusive in its totality for some time. Another point worth considering is captured in the following "prediction", see Schubert et al. (1979, p 170),

> "With regard to knowledge organization, we plan to continue the detailed development of generalization and topic hierarchies to determine how readily the full range of human concepts and human knowledge can be systematized in this way."

We started well to tackle this task, nonetheless, the "full range" has proven somewhat elusive to this point.

We pointed out the potential for learning within the topically oriented network organization and our desire to better understand our conception of the question-answering and problem-solving processes. This potential for learning and our better understanding remain to be fully exploited. Progress, however incremental, is steady and apparent and we expect to report additional successes in the years to come. Additional research needs to be undertaken before it will be possible to accurately access how daunting a task lies ahead in dealing with the overall general problems of knowledge representation and organization. A hopeful sign is the fact that at least over the past decade many additional researchers have generated a wealth of new research results in these areas and the computational paradigm is now being much more widely applied.

References

Abelson, R., The Structure of Belief Systems, in: R.Schank, and K. Colby, (Eds.), Computer Models of Thought and Language, (W.H. Freeman, San Francisco, Ca. 1973).

Anderson, J. and Bower, G., Human Associative Memory, (Winston and Sons, Wash., D.C., 1973).

Bartsch, R. and Vennemann, T., Semantic Structures, (Athenaum Verlag, Frankfurt, F.R.G., 1972).

Bates, M., The Theory and Practice of Augmented Transition Network Grammars. in: L. Bolc (Ed.), Natural Language Communications with Computers, (Springer-Verlag, Munich, 1978).

Bobrow, D., Dimensions of Representation. in: D. Bobrow and A. Collins (Eds.), Representation and Understanding, Academic Press, New York, 1975).

Bobrow, D., Special Issue on Non-Monotonic Logic of Artificial Intelligence Journal, (North Holland, Amsterdam, 1980).

Bobrow, D., & Winograd, T., An Overview of KRL: a Knowledge Representation Language, Cognitive Science 1(1) (1977) 3-46.

Bobrow, D., Kaplan, R., Kay, M., Norman, D., Thompson, H., and Winograd, T., GUS: a Frame Driven Dialog System. Artificial Intelligence 8(2) (1977) 155-174.

Brachman R., What's in a Concept: Structural Foundations for Semantic Networks, International Journal of Man-Machine Studies 9(2) (1977)127-152.

Brachman, R., On the Epistemological Status of Semantic Networks, in: N. Findler (Ed.), Associative Networks: The Representation and Use of Knowledge by Machine, (Academic Press, NY, 1979).

Brachman, R., and Smith, B., Special Issue on Knowledge Representation for SIGART, (ACM, New York, 1980).

Brachman, R. J., Gilbert, V. P., and Levesque, H. J. An Essential Hybrid Reasoning System: Knowledge and Symbol Level Accounts of Krypton. In Proceedings of the Ninth International Joint Conference on Artificial Intelligence, Los Angeles, California, 1985, 532-539.

Brady, J., and Wielinga, R., Reading the Writing on the Wall. in:

E. Riseman, and A. Hansen, (Eds.), Computer Vision Systems, (Academic Press, New York, 1978).
Brown, J. S., and Burton, R., Diagnostic Models for Procedural Bugs in Basic Mathematical Skills, Cognitive Science 2(2) (1978) 155-192.
Brown, J., Burton, R., and Bell, A., SOPHIE: A Sophisticated Instructional Environment for Teaching Electronic Troubleshooting (An Example of AI in CAI), Technical Report BBN Report No. 2790, Bolt Beranek and Newman, Inc.,1974.
Bruce, B., A Model for Temporal References and its Application in a Question-Answering Program, Artificial Intelligence 3(1) (1972) 1-26.
Bundy, A., Byrd, L., and Mellish, C., Special Purpose, but Domain Independent Inference Mechanisms. In Proc. ECAI-82, Orsay, France, 1982, 67-74.
Buchanan, B., Sutherland, G., and Feigenbaum, E., Heuristic DENDRAL: A Program for Generating Explanatory Hypothesis in Organic Chemistry, in: B. Meltzer and D. Michie (Eds.), Machine Intelligence 4, (Edinburgh University Press, Edinburgh, 1969).
Cercone, N., The Nature and Computational Use of a Meaning Representation for Word Concepts, American Journal for Computational Linguistics 2 (1975) 64-81.
Cercone, N., The Representation and Use of Knowledge in an Associative Network for the Automatic Comprehension of Natural Language, in: L. Bolc (Ed.), Natural Language Based Computer Systems, (C. Hanser Verlag,Muchen, West Germany, 1980).
Cercone, N., Hadley, R., Martin, F., McFetridge, P., and Strzalkowski, T., Designing and Automating the Quality Assessment of a Knowledge-Based System: The Initial Automated Academic Advisor Experience, IEEE Proceedings of the Workshop on Principles of Knowledge-Based Systems, Denver.
Cercone, N. and Schubert, L., Toward a State-Based Conceptual Representation, IJCAI75, Fourth International Joint Conference on Artificial Intelligence, Tbilisi, USSR, 1975, 83-91.
Collins, A., and Quillian, M., How to Make a Language User, in: E. Tulving and W. Donaldson (Eds.), Organisation of Memory, Academic Press, New York, 1972).
Cresswell, M., Logics and Languages, (Methuen,London, England, 1973).
de Haan, J. & Schubert, L. K. Inference in a topically organized semantic net. In Proceedings of the Fifth National Conference on Artificial Intelligence, pages 334--338, Philadelphia, PA., 1986.
Duda, R., Gashning, J., Hart, P., Konolige, K., Reboh, R., Barrett, P., and Slocum, J., Development of the PROSPECTOR Consultation System for Mineral Exploration, Final Report, SRI Projects 5821 and 6415. Technical Report 5821 & 6415, SRI International, 1978.
Fahlman, S., A System for Representing and Using Real World Knowledge, Technical Report AI Lab Memo 331, Project MAC, M.I.T., 1975.
Fahlman, S. NETL: A System for Representing and Using Real World Knowledge, (M.I.T. Press, Cambridge, Mass., 1979).
Fahlman, S. Three Flavors of Parallelism, 230-235. CSCSI 82, Proceedings of the Canadian Society for Computational Studies of Intelligence, Saskatoon, Sask., 1982.
Fillmore, C. J. The Case for Case. In Bach, E., and Harms, R. (Eds.), Universals in Linguistic Theory, New York, New York: Holt, Rinehart, and Winston, 1968.
Findler, N. (Ed.), Associative Networks: The Representation and Use of Knowledge by Machine, (Academic Press, New York, 1979).
Fodor, J. Troubles About Actions. In Davidson, D., and Harman, G. (Eds.), Semantics of Natural Language, Boston: D. Reidel Publishing Co., 1972.
Forgy, C., and McDermott, J. OPS, A Domain Independent Production System Language, 933-939. IJCAI77, Proceedings of the 5th International Joint Conference on Artificial Intelligence, Cambridge, Mass., 1977.
Goebel, R. Organising Factual Knowledge in a Semantic Network. Technical Report TR77-8, Department of Computing Science, University of Alberta, 1977.
Goldstein, I., and Roberts, R. NUDGE: A Knowledge-Based Scheduling Program, 257-263. IJCAI77, Proceedings of the 5th International Joint Conference on Artificial Intelligence, Cambridge, Mass., 1977.
Green, C. A Summary of the PSI Program Synthesis System, 380-381. IJCAI77, Proceedings of the 4th International Joint Conference on Artificial Intelligence, Cambridge, Massachusetts, 1977.
Guzman, A. Computer Recognition of Three Dimensional Objects in a Visual Scene. Technical Report MAC-TR-59, AI Laboratory, M.I.T., 1968.
Havens, W. A Procedural Model of Recognition, 263-264. Proceedings of the 5th IJCAI, MIT, Cambridge, MA, 1977.
Havens, W., and Mackworth, A. Representing Knowledge of theVisual World. In McCalla, G., and Cercone, N. (Eds.), IEEE Computer special issue on Knowledge Representation, New York: IEEE, 1983.
Hayes, P. Some Issues and Non Issues in Representation Theory, 63-79. Proceedings of the AISB Conference, University of Sussex, Sussex, England., 1974.
Hayes, P. In Defence of Logic, 559-565. IJCAI77, Proceedings of the 5th International Joint Conference on Artificial Intelligence, Cambridge, Mass., 1977.
Hayes-Roth, F., Waterman, D. A., and Lenat, D. B., (Editors). Building Expert Systems. Reading, Mass.:Addison-Wesley,1983.
Hendrix, G. Encoding Knowledge in Partitioned Networks. In N. Findler (Ed.), Associative Networks: The Representation and Use of Knowledge by Machine, New York: Academic Press,1979.
Hewitt, C. Description and Theoretical Analysis of PLANNER. Technical Report Ph.D. Thesis, AI Laboratory, M.I.T., 1972.
Jenkins, T., Gaillard, A., Holmback, H., Namioka, A., Darvish, J., Harrison, P., & Lorbeski, M. Automated message understanding: A real-world prototype. In Proceedings of the Third International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems (IEA/AIE), pages 546-552, Charleston, S.C., 1990.
Kay, M. Xerox's GUS (Genial Understanding System), 351-360. Proceedings Symposium on Advanced Memory Concepts, Menlo Park, Ca., 1976.
Kochut, K. Toward the Elastic ATN Implementation. In L. Bolc (Eds.), The Design of Interpreters, Compilers, and Editors for Augmented Transition Networks, Munich: Springer-Verlag,1983.
Lakoff, G. Linguistics and Natural Logic. In Davidson, D., and Harman, G. (Eds.), Semantics of Natural Language, Boston: D. Reidel Publishing Co., 1972.
Levesque, H. A Procedural Approach to Semantic Networks. Technical Report TR-105, Department of Computer Science, University of Toronto, 1977.
Levesque, H., and Mylopoulos, J. A Procedural Semantics for Semantic Networks. In N. Findler (Ed.), Associative Networks: The Representation and Use of Knowledge by Machine, New York: Academic Press, 1979.
Levesque, H. J. and Brachman, R. J., Expressiveness and Tractability in Knowledge Representation and Reasoning, Computational Intelligence 3 (1987) 78--93.
Lewis, D. General Semantics. In Davidson, D., and Harman, G. (Eds.), Semantics of Natural Language, Boston: D. Reidel Publishing Co., 1972.
Marcus, M. Theory of Syntactic Recognition for Natural Language. Cambridge, Mass.:MIT Press,1980.
Mathlab Group. MACSYMA Reference Manual. Cambridge, Massachusetts:Computer Science Laboratory, Massachusetts Institute of Technology, 1977.
McCalla, G. An Approach to the Organisation of Knowledge for the Modelling of Conversation. Technical Report PhD Thesis, University of British Columbia, 1977.
McCalla, G., and Cercone, N. Special Issue on Knowledge Representation of IEEE Computer. New York:IEEE, 1983.
McCarthy, J., and Hayes, P. Some Philosophical Problems from the Standpoint of Artificial Intelligence. In Meltzer, B., and Michie, D. (Eds.), Machine Intelligence, New York: American Elsevier, 1969.
McCarty, L. and Sridharan, N. The Representation of an Evolving System of Legal Concepts II: Prototypes and Deformations, 246-253. IJCAI81, Proceedings of the 7th International Joint Conference on Artificial Intelligence,Vancouver, Canada, 1981.
McDermott, D. A Packet-Based Approach to the Symbol Mapping

Problem. SIGART, 1975, (53)}, 6-7.
McSkimmin, J.R., and Minker, J. The Use of a Semantic Network in a Deductive Question-Answering System, 50-58. IJCAI77, Proceedings of the 5th International Joint Conference on Artificial Intelligence, Cambridge, Mass.,1977.
Miller, S. A. and Schubert, L. K., Time Revisited. In CSCSI-88, Edmonton, Alberta, 1988a.
Miller, S. A. and Schubert, L. K., Using Specialists to Accelerate General Reasoning. In Proceedings of the Seventh National Conference on Artificial Intelligence, 161-165, St Paul, Minnesota, 1988b.
Minsky, M. A Framework for Representing Knowledge. In P. Winston (Ed.), Psychology of Computer Vision, New York: McGraw Hill, 1975.
Montague, R. Pragmatics and Intensional Logic. In Davidson, D., and Harman, G. (Eds.), Semantics of Natural Language, Boston: D. Reidel Publishing Co., 1972.
Montague, R. The Proper Treatment of Quantification in Ordinary English. In R. Thomason (Ed.), Formal Philosophy: Selected Papers of Richard Montague, New Haven, Conn.: Yale University Press, 1974.
Montague, R. English as a Formal Language. In R. Thomason (Ed.), Formal Philosophy: Selected Papers of Richard Montague, New Haven, Conn.: Yale University Press, 1974.
Moore, R. D-SCRIPT: A Computational Theory of Descriptions. Technical Report Memo No. 278, Artificial Intelligence Lab, M.I.T., 1973.
Mylopoulos, J., Cohen, P., Borgida, A., and Sugar, L. Semantic Networks and the Generation of Context, 34-142. IJCAI75, Advance Papers of the 4th International Joint Conference on Artificial Intelligence, Tbilisi, USSR, 1975.
Mylopoulos, J., Shibahara, T., and Tsotsos, J. Building Knowledge-Based Systems: The PSN Experience. In McCalla, G., and Cercone, N. (Eds.), IEEE Computer special issue on Knowledge Representation, New York: IEEE, 1983.
Newell, A. Production Systems: Models of Control Structure. In W. Chase (Eds.), Visual Information Processing, New York: Academic Press, 1973.
Nii, H., and Aiello, N. AGE: A Knowledge-Based Program for Building Knowledge-Based Programs, 645-655. IJCAI79, Proceedings of the 6th International Joint Conference on Artificial Intelligence, Tokyo, Japan, 1979.
Norman, D., and Rumelhart, D. Explorations in Cognition. San Francisco, Ca.:W.H. Freeman, 1975.
Pople, H. The Formation of Composite Hypotheses in Diagnostic Problem Solving - An Exercise in Synthetic Reasoning INTERNIST], 1030-1037. IJCAI77, Proceedings of the 5th International Joint Conference on Artificial Intelligence, Cambridge, Massachusetts, 1977.
Quillian, M. Semantic Memory. In Minsky, M. (Ed.), Semantic Information Processing, Cambridge, Ma.: M.I.T. Press, 1968.
Quillian, M. The Teachable Language Comprehender. Comm. of the ACM, 1969, 12(8), 459-476.
Quine, W. Word and Object. Cambridge, Ma.:M.I.T. Press, 1960.
Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. A Grammar of Contempory English. London:Seminar Press, 1972.
Rich, C. TheLayered Architecture of a System for Reasoning About Programs. In Proceedings of the 9th International Joint Conference on Artificial Intelligence, Los Angeles, California, 540-546, 1985.
Roberts, R. and Goldstein, I. The FRL Primer. Technical Report AI Memo 408, MIT, November 1977.
Rumelhart, D., Lindsay, P., and Norman, D. A Process Model for Long Term Memory. In Tulving, E., and Donaldson, D. (Ed.), Organisation of Memory, New York: Academic Press, 1972.
Schank, R. Conceptual Dependency: A Theory of Natural Language Understanding. Cognitive Psychology, 1972, 3, 552-631.
Schank, R. Adverbs of Belief. Lingua, 1974, 33, 45-67.
Schank, R. The Role of Memory in Language Processing. In C. Cofer (Ed.), The Structure of Human Memory, San Francisco: Freeman, 1975.
Schank, R., and Abelson, R. Scripts, Plans, and Knowledge,151-157. IJCAI75, Advance Papers of the 4th International Joint Conference on Artificial Intelligence, Tbilisi, USSR,1975.
Schank, R., Goldman, N., Rieger, C., and Riesbeck, C. Margie: Memory, Analysis, Response Generation and Inference on English, 255-261. IJCAI73, Proceedings of the 3rd International Joint Conference on Artificial Intelligence, Stanford, Ca., 1973.
Schubert, L. Extending the Expressive Power of Semantic Networks, 158-164. IJCAI75, Advance Papers of the 4th International Joint Conference on Artificial Intelligence, Tbilisi, USSR, 1975.
Schubert, L. Extending the Expressive Power of Semantic Networks. Artificial Intelligence, 1976, 7, 163-198.
Schubert, L., Cercone, N., and Goebel, R. The Structure and Organisation of a Semantic Network for Comprehension and Inference. Technical Report TR78-1, Department of Computing Science, University of Alberta, 1978.
Schubert, L., Goebel, R., and Cercone, N. The Structure and Organisation of a Semantic Network for Comprehension and Inference. In N. Findler (Ed.), Associative Networks: The Representation and Use of Knowledge by Machine, New York: Academic Press, 1979.
Schubert, L. K., Problems with Parts. In Proceedings of the Sixth International Joint Conference on Artificial Intelligence, Tokyo, Japan, 778-784, 1979.
Schubert, L. K., Papalaskaris, M. A., and Taugher, J., Accelerating Deductive Inference: Special Methods for Taxonomies, Colours and Times. In Cercone, N. and McCalla, G., editors, The Knowledge Frontier, New York: Springer-Verlag, 1987.
Schubert, L. K. & Hwang, C. H. A logical approach to goal-based and causal inferences in narrative understanding. Technical Report BCS-G2010-72, The Boeing Company, Seattle,1988.
Schubert, L. K. & Hwang, C. H. An episodic knowledge representation for narrative texts. In Brachman, R. J., Levesque, H. J., and Reiter, R., editors, Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning, pages 444--458, Toronto, Ontario. Morgan-Kaufmann, 1989.
Schubert, L. K., Schaeffer, S., & Hwang, C. H. EPILOG: The computational system for Episodic Logic. Prepared for Boeing Co. under Purchase Contract W-278258, 1990.
Scragg, G. Ph.D. Thesis. Technical Report, Department of Computing Science, University of California, San Diego,1975.
Shapiro, S. A Net Structure for Semantic Information Storage, Deduction, and Retrieval, 512-523. IJCAI71, Proceedings of the 2nd International Joint Conference on Artificial Intelligence, London, England, 1971.
Shapiro, S. The SNePS Semantic Network Processing System. In N. Findler (Ed.), Associative Networks: The Representation and Use of Knowledge by Machine, New York: Academic Press,1979.
Shortliffe, E. Computer - Based Medical Consultation: MYCIN. New York:North-Holland, 1976.
Simmons, R. Semantic Networks: Their Computation and Use for Understanding English Sentences. In Schank, R., and Colby, K. (Eds.), Computer Models of Thought and Language, SanFrancisco, Ca.: W.H. Freeman, 1973.
Slagle, J. R., Automatic Theorem Proving with Built-In Theories Including Equality, Partial Ordering, and Sets. Journal of the ACM 19(1):120-135, 1972.
Stickel, M. E. (1985). Automated Deduction by Theory Resolution. Journal of Automated Reasoning 1, 333-335, 1985.
Strzalkowski, T. ENGRA-Yet Another Parser for English. Technical Report TR83-10, Computing Science Dept., Simon Fraser University, 1983.
Strzalkowski, T., and Cercone, N. A Framework for Computing Extra Sentential References. Computational Intelligence 2(4), 1986, 159-180.
Strzalkowski, T. and Cercone, N. Non Singular Concepts in Natural Language Discourse. Computational Linguistics 15(4), 1989, 171-187.
van Melle, W. A Domain Independent System that Aids in Constructing Consultation Programs (EMYCIN). Technical Report STAN-CS-80-820, Department of Computer Science, Stanford University, 1980.
Vilain, M., The Restricted Language Architecture of a Hybrid Representation System. In Proceedings of the Ninth International Joint Conference on Artificial Intelligence, Los Angeles, California, 547-551, 1985.
Vogel, C., Cercone, N., and Popowich, F. Practical Aspects of Inheritance Structures, Technical Report, Intelligent Systems

Group, Centre for Systems Science, Simon Fraser University, Burnaby, B.C., 1990.
Weiss, S. M., and Kulikowski, C. A. A Practical Guide to Designing Expert Systems. Totowa, N. J.:Rowman and Allanheld, 1984.
Wilks, Y. Preference Semantics. Technical Report AIM-206,Stanford AI Project, Stanford University, 1973.
Wilks, Y. Primitives and Words, 34-37. Advance Papers of Theoretical Issues in Natural Language Processing Workshop, Champaign-Urbana, Ill., 1975.
Wilks, Y. Good and Bad Arguments about Semantic Primitives. Technical Report DAI-42, University of Edinburgh, 1977.
Wilson, K. From Association to Structure. Amsterdam:North Holland, 1979.
Winograd, T. Language as a Cognitive Process. Reading, Mass.:Addison-Wesley, 1983.
Winston, P. Learning Structural Descriptions from Examples. Technical Report MAC-TR-76, PRoject MAC, M.I.T., 1970.
Woods, W. Transition Networks for Natural Language Analysis. Communications of the ACM, 1970, 13, 591-606.
Woods, W. What's in a Link. In Bobrow, D., and Collins, A. (Eds.), Representation and Understanding, New York: Academic Press, 1975.
Woods, W. What's Important about Knowledge Representation. In McCalla, G., and Cercone, N. (Eds.), IEEE Computer special issue on Knowledge Representation, New York: IEEE, 1983.
Zadeh, L. A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges. Journal of Cybernetics, 1972, 2, 4-34.