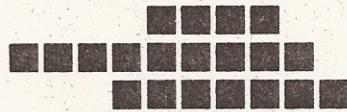


# NAIC

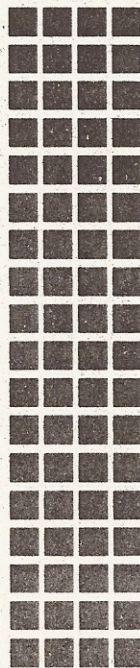


NORTHEAST  
ARTIFICIAL INTELLIGENCE  
CONSORTIUM



1987 Natural  
Language Planning  
Workshop

# TECHNICAL REPORT SERIES



Presentations from the

# 1987 Natural Language Planning Workshop

"PLANNING FOR FUTURE RESEARCH:  
DIRECTIONS FOR THE NEXT DECADE"

**Sponsored by:**  
Northeast Artificial Intelligence Consortium

September 20-23, 1987  
*Minnowbrook Conference Center*  
*Blue Mountain Lake, NY*

Northeast Artificial Intelligence Consortium  
WR-8703

This work was supported by the Air Force Systems Command, Rome Air Development Center, Griffiss Air Force Base, New York 13441-5700, and the Air Force Office of Scientific Research, Bolling AFB, DC 20332 under Contract No. F30602-85-C-0008. This contract supports the Northeast Artificial Intelligence Consortium (NAIC).

# KNOWLEDGE REPRESENTATION FOR NATURAL LANGUAGE PROCESSING<sup>1</sup>

Stuart C. Shapiro<sup>2</sup> and William J. Rapaport

Department of Computer Science  
State University of New York at Buffalo  
Buffalo, NY 14260  
(716) 636-3182

shapiro@gort.cs.buffalo.edu, rapaport@gort.cs.buffalo.edu

## ABSTRACT

We extend, deepen, and clarify our theory of intensional knowledge representation for natural-language processing, as presented in previous papers and in light of objections raised by others. The essential claim is that tokens in a knowledge-representation system represent only intensions and not extensions. We are pursuing this investigation by building CASSIE, a computer model of a cognitive agent and, to the extent she works, a cognitive agent herself. CASSIE's mind is implemented in the SNePS propositional semantic-network processing system. In this paper, we explicate the relations among nodes, mental tokens, concepts, actual objects, concepts in the belief spaces of different cognitive agents, concepts in the belief spaces of an agent and the agent's model of other agents, concepts of other cognitive agents, propositions, and concepts of concepts.

## 1. INTRODUCTION

Knowledge representation (KR) is concerned, among other things, with the representation of information in artificial-intelligence (AI) computer systems. Few, if any, researchers claim at the outset to be working on an AI system whose abilities are co-extensive with human intellectual abilities, so most AI systems are designed to work on a narrower (if still broad) domain, such as vision, problem solving, robotics, or natural-language processing (NLP). It may be argued that in order to reach human-level abilities in any one of these areas, it is necessary to have human-level abilities in all the others. If that is true, then choosing an area of AI to work on is just a research strategy for working on general AI, but in any case the choice determines what one thinks are the problems to be worked on (initially) and the information that needs to be represented. In this paper, we shall consider the information that needs to be represented by an AI system whose domain is general NLP. By natural-language processing, we mean the use of natural human languages (e.g., English) for the kind of communication with agents (human or computer) that humans engage in when they communicate with each other in those languages. In particular, we exclude the kinds of text processing that treat natural-language (NL) texts as uninterpreted strings of characters (such as producing a concordance), and we include both NL generation and NL understanding. The quintessential NLP task is interactive dialogue, although other tasks (such as reading and summarizing a narrative and translating a text from one NL to another) have been recognized by AI researchers as requiring the same abilities.

In order to focus on the information that an NLP system needs, it is useful to consider briefly, by way of contrast, the information needed by some other intelligent systems (remembering, however, that as an NLP system is extended into a general intelligent system, it will need such information

<sup>1</sup> This paper is based in part upon work supported by the National Science Foundation under Grant Nos. IST-8504713 (Rapaport) and IRI-8610517 (Rapaport and Shapiro), and in part by the Air Force Systems Command, Rome Air Development Center, Griffiss Air Force Base, New York 13441-5700, and the Air Force Office of Scientific Research, Bolling AFB DC 20332 under contract No. F30602-85-C-0008 to the Northeast Artificial Intelligence Consortium (Shapiro).

<sup>2</sup> During 1987-88 on sabbatical at USC/Information Sciences Institute, 4676 Admiralty Way, Marina del Rey, CA 90292-6695

also). Robot systems (both for locomotion and for manipulation) and vision systems must operate in the here and now of the real world. The way the world really is (the subject of physics) is important to them, because they must operate in it. They must recognize multiple interactions with the same physical object as *being* interactions with the same physical object. This is the essence of the stereo vision problem: a mobile robot can go between two objects but must go around one; a manipulator needs at least two points of contact to grasp an object. On the other hand, although face-to-face dialogue often includes references to objects in the current environment, the essence of NL is the ability to discuss the not-here and the not-now.

Distinguishing "knowledge representation" from "data storage" is not easy, but Brian Cantwell Smith's Knowledge-Representation Hypothesis is fairly well-accepted:

Any mechanically embodied intelligent process will be comprised of structural ingredients that a) we as external observers naturally take to represent a propositional account of the knowledge that the overall process exhibits, and b) independent of such external semantical attribution, play a formal but causal and essential role in engendering the behaviour that manifests that knowledge. (Smith 1982: 33.)

As is often the case in AI, this takes key terms (such as "knowledge") on their pre-theoretic meaning, to be attributed (or not) to AI systems by the same criteria by which they are attributed to humans.

The declarative/procedural controversy (Winograd 1975) has fairly well died out, and it would seem that the "structural ingredients" in Smith's KR Hypothesis, which are to be taken as the represented knowledge, could be either declarative or procedural. It is clear that the declarative knowledge of a system could be considered procedural to the extent that it plays a "causal and essential role in engendering ... behaviour". However, the distinction is still useful if worded in a slightly different way than it has been in the past. Let us consider as *procedural knowledge* that knowledge that *only plays such a causal role, and* let us consider as *declarative knowledge* that knowledge that the system can discuss with another. So, for example, humans act as though they know the grammar of their native language (competence), but most cannot describe that grammar at all, and no one has yet been able to describe it completely. So we would say that people have procedural knowledge of the grammar of their native language, but not declarative knowledge of it. In this paper, we shall concentrate on discussing the declarative knowledge needed by a system for human-level NLP abilities, rather than the procedural knowledge needed for NLP. (For a discussion of the latter, see Weischedel 1986. Other recent reviews of KR are Barr & Davidson 1981, Mylopoulos & Levesque 1984, Levesque 1986, and Kramer & Mylopoulos 1987. Recent collections of papers on KR relevant to NLP include Brachman & Levesque 1985, King & Rosner 1986, and McCalla & Cercone 1987.)

## 2. INTENSIONAL KNOWLEDGE REPRESENTATION

Insofar as an AI NLP system is considered a model of a cognitive agent, the information represented in its "mind" consists of the beliefs, knowledge, and other intentional (i.e., psychological) attitudes of the agent (together with the objects, properties, and relations that those attitudes are directed to) (cf. Brachman & Smith 1980). Such a system must also be able to reason about the beliefs, knowledge, and other intentional attitudes of itself and of other cognitive agents with which it interacts.

In a series of earlier papers, we have argued that such a system must represent and reason about *intensional*, as opposed to *extensional*, entities (Maida & Shapiro 1982; Rapaport & Shapiro 1984; Rapaport 1985, 1986a; Wiebe & Rapaport 1986; Shapiro & Rapaport 1986, 1987). Briefly, this means that, instead of storing information about *objects in the world*, as vision and robot systems must, NLP systems store information about *mental objects* that may or may not even exist. In the main body of this paper, we will survey the kinds of mental objects NLP systems must deal with.

### 3. KR FORMALISMS.

#### 3.1. Syntax and Semantics.

A KR formalism, like any language, has a syntax and a semantics. The syntax specifies those parts of the language (such as "terms" and "sentences") that can be given meanings (interpretations) by the semantics. Larger syntactic parts (such as sentences) are formed from smaller syntactic parts by putting the smaller parts together in various structured ways with the use of *punctuation*. We presume that the reader is already familiar with this, but, as a reminder, here is a typical example for a first-order language  $\mathcal{L}$  (adapted from Rapaport 1987):

#### SYNTAX OF $\mathcal{L}$ :

##### Alphabet:

<i>n</i> -place predicate symbols:	$A, \dots, Z$ (with or without subscripts);
<i>n</i> -place function symbols:	$f, g, h$ (with or without subscripts);
individual variables:	$u, \dots, z$ (with or without subscripts);
individual constants:	$a, \dots, e$ (with or without subscripts);
connectives:	$\neg, \vee, \wedge, \rightarrow$
punctuation:	$',' , '[' , ']' , '(' , ')'$
quantifiers:	$\forall, \exists$

##### Terms:

- (T1) All individual variables are *terms*.
- (T2) All individual constants are *terms*.
- (T3) If  $t_1, \dots, t_n$  are terms and  $f$  is an  $n$ -place function symbol, then  $f(t_1, \dots, t_n)$  is a *term*.
- (T4) Nothing else is a term.

##### Well-formed formulas:

- (WFF.1) If  $t_1, \dots, t_n$  are terms and  $P$  is an  $n$ -place predicate symbol, then  $P(t_1, \dots, t_n)$  is a(n *atomic*) *well-formed formula*.
- (WFF.2) If  $\varphi$  and  $\psi$  are well-formed formulas,  $v$  is an individual variable, and ' $\varphi(v^*)$ ' is a well-formed formula containing zero or more occurrences of  $v$ , then  $\neg \varphi, (\varphi \vee \psi), (\varphi \wedge \psi), (\varphi \rightarrow \psi), \forall v[\varphi(v^*)], \exists v[\varphi(v^*)]$  are *well-formed formulas*.
- (WFF.3) Nothing else is a well-formed formula.

#### SEMANTICS OF $\mathcal{L}$ :

Let  $M$  be the structure  $\langle D, R, F \rangle$ , where  $D$  is a non-empty set,  $R$  is a set of  $n$ -place relations on the elements of  $D$ , and  $F$  is a set of  $n$ -place functions on the elements of  $D$ . An *interpretation*,  $I$ , on  $M$  for  $\mathcal{L}$  is a function from the symbols of  $\mathcal{L}$  to  $D \cup R \cup F$  such that:

- (1) If  $t$  is an individual constant or individual variable, then  $I(t) \in D$ .
- (2) If  $f$  is a function symbol, then  $I(f) \in F$ .
- (3) If  $f$  is an  $n$ -place function symbol and  $t_1, \dots, t_n$  are terms, then  $I(f(t_1, \dots, t_n)) = I(f)\langle I(t_1), \dots, I(t_n) \rangle \in D$ .
- (4) If  $P$  is an  $n$ -place predicate symbol, then  $I(P) \in R$ .

The notion of "truth on an interpretation" (symbolized as:  $\models_I$ ) can be defined recursively as follows:

- (1) If  $P$  is an  $n$ -place predicate symbol, and  $t_1, \dots, t_n$  are terms, then  $\models_I P(t_1, \dots, t_n)$  if and only if  $\langle I(t_1), \dots, I(t_n) \rangle \in I(P)$ .
- (2) If  $\varphi$  and  $\psi$  are WFFs and  $v$  is an individual variable, then
  - (a)  $\models_I \neg \varphi$  if and only if not  $\models_I \varphi$ ;

- (b)  $\models_I(\varphi \vee \psi)$  if and only if  $\models_I \varphi$  or  $\models_I \psi$ ;
- (c)  $\models_I(\varphi \wedge \psi)$  if and only if  $\models_I \varphi$  and  $\models_I \psi$ ;
- (d)  $\models_I(\varphi \rightarrow \psi)$  if and only if not  $\models_I \varphi$  or  $\models_I \psi$ ;
- (e)  $\models_I \forall v \varphi$  if and only if  $\models_{I'} \varphi$  for every interpretation  $I'$  that differs from  $I$  at most on what it assigns to  $v$ ;
- (f)  $\models_I \exists v \varphi$  if and only if  $\models_{I'} \varphi$  for some interpretation  $I'$  that differs from  $I$  at most on what it assigns to  $v$ .

Finally,

- (3) A WFF  $\varphi$  is *valid in  $M$*  (written:  $M \models \varphi$ ) if and only if  $\models_I \varphi$  for every interpretation  $I$  on  $M$ .
- (4) A structure  $M$  is a *model* for a set  $H$  of WFFs if and only if  $M \models H_i$  for every WFF  $H_i \in H$ .

Finally, one often says that the interpretation of a formula is a truth value, either **T** or **F**.

Let's note several things about these definitions. First, there is a distinction between terms and formulas. The interpretations of terms are elements of the domain  $D$ , but interpretations of formulas are elements of the set  $\{\mathbf{T}, \mathbf{F}\}$ ; variables vary over elements of  $D$ . These distinctions between terms and formulas, together with the fact that terms can be quantified over—i.e., talked about—are what makes this language *first-order*. Secondly, the punctuation consists of marks like commas and parentheses, and of linear order. For example, if  $f$  is a two-place function symbol and  $a$  and  $b$  are constants, then  $f(a, b)$  and  $f(b, a)$  are different terms. The well-known propositional connectives are also punctuation marks. Punctuation marks do not receive interpretations; only constants, function symbols, predicate symbols, and formulas do. Only constants and functional terms containing no variables have interpretations that are in the domain  $D$  (unless **T** and **F** are, themselves, in  $D$ ; this is not ruled out, but is very uncommon). The domain  $D$ , itself, is unspecified—it can be anything. Similarly unspecified are the mappings that are the interpretations of the function and predicate symbols. The main thing specified by the semantics is the way the truth value of a formula is calculated from the truth values of its atomic parts.

In a language such as we have shown here, the formulas are the sentences—what can be said in the language. The elements of the domain of interpretation are the individuals that can be talked about—that can be discussed. Among the things that *cannot* be discussed in such a language are the punctuation marks, the mappings, the symbols, the formulas, and what the formulas say (not only **T** and **F**, but what distinguishes  $(P \rightarrow P)$  from  $(Q \vee \neg Q)$ ).

Every KR formalism has a more or less well-defined syntax, semantics, and intended domain of interpretation. These determine the individuals that can be discussed in the formalism and what can be said about them.

## 3.2. SEMANTICS.

### 3.2.1. The Intended Domain.

Above, we said that, instead of storing information about objects in the world, as vision and robot systems must, NLP systems store information about mental objects that may or may not even exist. Thus, the intended domain of interpretation of NLP systems is not the real world of real objects, but a mental world of (possibly) imaginary objects—“*Aussersein*”: the domain of the objects of thought—the things we can think about, have beliefs about, etc., whether or not they exist or are true (Meinong 1904; Rapaport 1978, 1985; Shapiro & Rapaport 1986, 1987). These elements can be called “mental objects” or “intensions”. They are what we refer to as “concepts”.

What is the relation between a concept (an intensional object in *Aussersein*) and objects in the world? Some concepts have objects in the world that *exemplify* them (in Rapaport 1978, these objects are referred to as the *Sein-correlates* of the concepts); some have one; some have more than one. Some have none: for instance, concepts of unicorns and square circles have no objects exemplifying them in

the real world. The concept of the coffee cup that I<sup>3</sup> habitually use in and around my office has one object that exemplifies it. The concept of the straight-backed teak chair with black seat that I<sup>4</sup> have in my house has eight objects that exemplify it.

I<sup>5</sup> believe that I can recognize my coffee cup, but it may be that there is one exactly like it. If, some night, someone were to replace my coffee cup with one that looked exactly like it, I might not notice the difference (cf. *The Ballad of Shakey's Pizza Parlor*, Dennett 1982: 53-60). In that case I would take the new cup to be my cup, the object that exemplifies the concept denoted by the mental token I have been using for that concept for years, and I would not create a new mental token for it. On the other hand, I know I cannot distinguish my eight chairs at home, and have no beliefs about any individual one.

What do I mean by "my concept of X", and how does that relate to "your concept of X"? If my beliefs about X coincide exactly with your beliefs about X (e.g., no matter how long we discuss X, we find no disagreement), then our concepts of X are the same intensional objects (although there would be two symbols representing them, one in my mind, one in yours). However, if we disagree in the slightest about X, then our concepts of X are different intensional objects in *Aussersein*, even if they are exemplified by one and the same object in the real world.

### 3.2.2. Epistemological Ontology.

What sorts of things are there in *Aussersein*? Since this is a question of what there is among mental objects, it is a question of epistemological ontology (cf. Rapaport 1985/1986). It can be answered by considering the naive ontology of things talked about in natural language (see Rapaport 1981, Hobbs 1985).

One sort of mental object is people. I have beliefs about Ronald Reagan; so I have a concept of Ronald Reagan. My concept of Ronald Reagan includes beliefs that he is currently President of the United States, was Governor of California, before that was an actor, etc. There is an object in the real world (at least, I believe so) that exemplifies the concept of a person named Ronald Reagan having these three properties. Depending on the rest of my beliefs about Ronald Reagan, there may or may not be an object in the real world exemplifying it. (Cf. Rapaport 1978.)

Another sort of mental object is mental acts, such as acts of believing. My concept of believing is being discussed in this paper, and there is a vast literature discussing other people's concepts of believing. Whether one believes that there is an object in the real world exemplifying believing depends on one's ontology (i.e., one's *non-epistemological ontology*), and the two authors of this paper sometimes disagree on that.

Another sort of mental object is propositions. Propositions are sometimes called "beliefs", "facts", or "truths", depending on what other beliefs the speaker has about the proposition being discussed. The ontological status of objects exemplifying propositions is as controversial as that of objects exemplifying beliefs.

There are other sorts of mental objects. Some are: properties, sentences, numbers, numerals, and truth values. These are all different sorts in the sense that we have different kinds of beliefs about them, and believe that different kinds of properties are relevant to them.

---

<sup>3</sup>When 'I' is used, the reader is invited to use a new mental token for a concept that is co-extensional with the reader's concept of one of the two authors, but the reader does not know which.

<sup>4</sup>The reader may identify the referent of this 'I' with the previous one, or may create a new mental token, again co-extensional with one of the two authors. In the latter case, the two 'I's might or might not refer to the same author, but in the former case, the reader is deciding that they refer to the same one, but still has not committed to which one.

<sup>5</sup>This now, surely, refers to the same author as one of the two previous 'I's, and the reader should have no trouble deciding which.

When we discuss people, we refer to them by using a proper name, a definite noun phrase, or, sometimes, an indefinite noun phrase. We can refer to Ronald Reagan as "Ronald Reagan" or "Nancy Reagan's favorite actor". A particular understander at a particular time will take these referring expressions to refer to his or her own concept of Ronald Reagan, and may react accordingly.

When we discuss propositions, we often use the phraseology, "that <clause>". For example, in the U.S. Declaration of Independence, we find, "We hold these Truths to be self-evident, that all Men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty, and the Pursuit of Happiness," and in Lincoln's Gettysburg Address we find, "a new nation, conceived in liberty and dedicated to the proposition that all men are created equal." We could also refer to a proposition using a definite noun phrase, e.g., "the first self-evident truth mentioned in the Declaration of Independence" or "Lincoln's favorite proposition". Whether the referent of "the first self-evident truth mentioned in the Declaration of Independence" *is the same as* the referent of "Lincoln's favorite proposition" depends on the particular beliefs of a particular understander.

People, believings, and propositions (those people, believings, and propositions we have beliefs about) are denizens of the same universe—Aussersein. If some cognitive agent believes that Lincoln believed that all men are created equal, part of that agent's Lincoln concept is that he participates in a believing act (its concept of believing) directed to the proposition that all men are created equal (its concept of that proposition).

#### 4. BRIEF SURVEY OF KR FORMALISMS.

Several different KR formalisms have been suggested for NLP systems. Here, we will survey a few of the major varieties, discussing their syntactic constructs and their intended domains of interpretation.

Conceptual Dependency theory (Schank & Rieger 1974, Schank 1975, Hardt 1987) uses a KR formalism consisting of sentences, called "conceptualizations", which assert the occurrence of events or states, and six types of terms: PPs—"real-world objects", ACTs—"real-world actions", PAs—"attributes of objects", AAs—"attributes of actions", Ts—"times", and LOCs—"locations". (The glosses of these types of terms are quoted from Schank & Rieger 1974: 378-379.) The set of ACTs is closed and consists of the well-known primitive ACTs PTRANS, ATRANS, etc. The syntax of an event conceptualization is a structure with six slots (or arguments), some of which are optional: actor, action, object, source, destination, and instrument. A stative conceptualization is a structure with an object, a state, and a value. Only certain types of terms can fill certain slots. For example, only a PP can be an actor, and only an ACT can be an action. Interestingly, conceptualizations, themselves, can be terms, although they are not one of the six official terms. For example, only a conceptualization can fill the instrument slot, a conceptualization can be located (MLOC) somewhere, and a conceptualization can cause another. A causation is another kind of conceptualization, consisting only of two slots, one containing the causing conceptualization and the other containing the caused conceptualization. Although, from the glosses of PP and ACT, it would seem that the intended domain of interpretation is the real world, the domain also must contain theoretically postulated objects such as: the "conscious processor" (CP) of people, in which conceptualizations are located; conditional events; and even negated events, which haven't happened.

KL-ONE (Brachman & Schmolze 1985) and its descendents, KL-TWO (Vilain 1985), KRYPTON (Brachman, Fikes, & Levesque 1983), and Loom (MacGregor & Bates 1987) separate their formalisms into two sub-languages—the definitional (or terminological) component, called the "TBox", and the assertional component, called the "ABox" (KL-ONE's separation is more rudimentary than that of the later systems). The ABox consists of sentences in a restricted first-order logic (e.g., KL-TWO doesn't allow any quantifiers) and are taken to assert truths in the domain of interest. Terms in the ABox have existential import. I.e., constants represent individuals that exist in the domain, and an existentially quantified sentence of the form  $\exists xP(x)$  asserts that an individual satisfying  $P$  exists in the domain. Unary predicate symbols, relational predicate symbols, and function symbols in ABox sentences are themselves terms in the TBox. There are two types of terms in the TBox—concepts and relations (or roles). TBox sentences make assertions like, "A family is: a kind of social-structure



with exactly 1 *male-parent* and this *male-parent* is a **man**; and a kind of **social-structure** with exactly 1 *female-parent* and this *female-parent* is a **woman**; and a kind of **social-structure** all of whose *children* are **persons**." Here, **family**, **social-structure**, **man**, **woman**, and **person** are concept terms; *male-parent*, *female-parent*, and *child* are relation terms; all other parts of the sentence are punctuation; and the sentence, itself, is taken as the definition of **family**. This last is most important—the sentence cannot be either true or false; it is simply a definition. If you think that this definition is incorrect, you cannot say, "No, you're wrong; a family isn't like that." You can only say, "Well, your concept of **family** is different from mine." Also, defining **family** in the TBox doesn't say that there are any in the domain of interest. That can only be said in the ABox by an existential statement such as  $\exists x \text{family}(x)$  or by an ABox assertion such as **family**(Shapiro). TBox sentences about relations are also definitional. For example, we could use the relation *child* to define *son* as a *child* that is a **man**. NIKL (Moser 1983) is another descendent of KL-ONE and consists only of a TBox (in fact, it is the TBox of KL-TWO), so all sentences in NIKL are definitional.

## 5. SNePS

One particularly well defined KR formalism for our purposes is SNePS/CASSIE. This is a particular application of SNePS, the Semantic Network Processing System (Shapiro 1979), with a particular set of arc labels and case frames, to model the mind of a cognitive agent referred to as CASSIE (the Cognitive Agent of the SNePS System—an Intelligent Entity; cf. Shapiro & Rapaport 1986, 1987).

### 5.1. Brief Overview of SNePS.

SNePS is a *propositional* semantic-network knowledge-representation and reasoning system. The following conditions hold in SNePS:

- (1) Each node represents a unique concept.
- (2) Each concept represented in the network is represented by a node.
- (3) Uniqueness Principle: Each concept represented in the network is represented by a unique node.
- (4) Arcs represent non-conceptual, binary, structural relations between nodes. (They are punctuation.)
- (5) The information represented about each concept is represented by the structure of the entire network connected to the node representing the concept.

Nodes represent only intensions and not extensions. Further, if we view SNePS as a system for modeling the mind of a cognitive agent, CASSIE, then all represented concepts are in CASSIE's mind. Thus, when CASSIE represents the beliefs of someone else, what is represented is CASSIE's representation of the beliefs of the other person, not the other person's actual beliefs.

### 5.2. Syntax and Semantics of SNePS/CASSIE.

The Uniqueness Principle guarantees that nodes represent intensional objects and that nodes will be shared whenever possible. Nodes that only have arcs pointing *to* them are considered to be unstructured or *atomic*. They include: (1) *sensory* nodes, which—when SNePS is being used to model a mind—represent interfaces with the external world (e.g., utterances); (2) *base* nodes, which represent individual concepts and properties; and (3) *variable* nodes, which represent arbitrary individuals (Fine 1983) or arbitrary propositions.

Molecular nodes, which have arcs emanating *from* them, include: (1) *structured individual* nodes, which represent structured individual concepts or properties (i.e., concepts and properties represented in such a way that their internal structure is exhibited) and (2) *structured proposition* nodes, which represent propositions; those with no incoming arcs represent CASSIE's *beliefs*. (Note that structured proposition nodes can also be considered to be structured individuals.) Proposition nodes are either *atomic* (representing atomic propositions) or are *rule nodes*. Rule nodes represent deduction rules and are used for node-based deductive inference (Shapiro 1978; Shapiro & McKay 1980; McKay &

Shapiro 1981; Shapiro, Martins, & McKay 1982). For each of the three categories of molecular nodes (structured individuals, atomic propositions, and rules), there are *constant* nodes of that category and *pattern* nodes of that category representing arbitrary entities of that category.

In this section, we give the syntax and semantics of some of the nodes and arcs used in interaction with CASSIE. We begin with a few rough definitions (cf. Shapiro 1979, Sect. 2.1, for more precise ones).

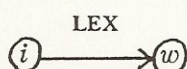
(Def. 1) A node *dominates* another node if there is a path of directed arcs from the first node to the second node.

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

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

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

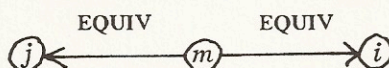
(Syn.1) If "*w*" is a(n English) word and "*i*" is an identifier not previously used, then



is a network, *w* is a sensory node, and *i* is a structured individual node.

(Sem.1) *i* is the objectum corresponding to the utterance of *w*.

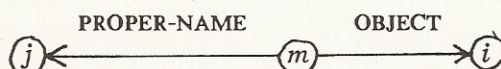
(Syn.2) If *i* and *j* are individual nodes, and "*m*" is an identifier not previously used, then



is a network and *m* is a structured proposition node.

(Sem.2) *m* is the objective corresponding to the proposition that objecta *i* and *j* (are believed by CASSIE to) correspond to the same actual object.

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

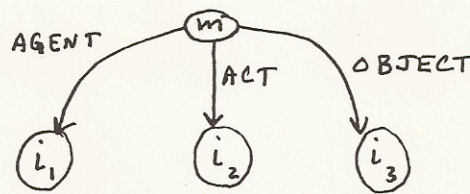


is a network and *m* is a structured proposition node.

(Sem.3) *m* is the objective corresponding to the proposition that objectum *i*'s proper name is *j*. (*j* is the objectum that is *i*'s proper name; its expression in English is represented by a node at the head of a LEX-arc emanating from *j*.)

(Syn.4) If  $i_1, i_2, i_3$  are individual nodes, previously used, then

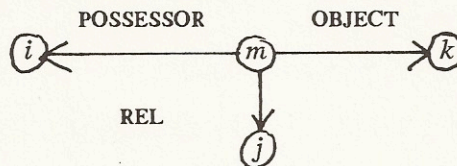
and "m" is an identifier not pre-



is a network and  $m$  is a structured proposition node. (In Figure 1, node p7 is an example of this for the mental act of believing; cf. Rapaport 1986a.)

(Sem.4)  $m$  is the objective corresponding to the proposition that agent  $i_1$  performs act  $i_2$  with respect to  $i_3$ .

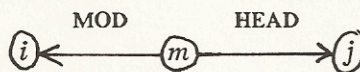
(Syn.5) If  $i, j$ , and  $k$  are individual nodes and "m" is an identifier not previously used, then



is a network and  $m$  is a structured proposition node.

(Sem.5)  $m$  is the objective corresponding to the proposition that  $k$  is  $i$ 's  $j$ .

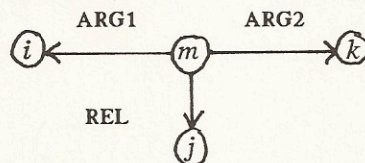
(Syn.6) If  $i$ , and  $j$  are individual nodes and "m" is an identifier not previously used, then



is a network and  $m$  is a structured individual node.

(Sem.6)  $m$  is the object corresponding the individual that is a  $j$  that is characterized as being  $i$ -ish.

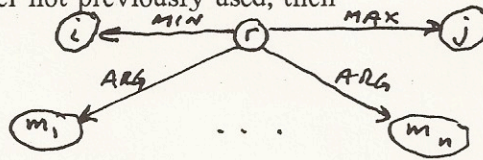
(Syn.7) If  $i, j$ , and  $k$  are individual nodes and "m" is an identifier not previously used, then



is a network and  $m$  is a structured proposition node.

(Sem.7)  $m$  is the objective corresponding to the proposition that  $i$  stands in relation  $j$  to  $k$ .

(Syn.8) If  $m_1, \dots, m_n$  are proposition nodes ( $n \geq 0$ ), “ $i$ ” and “ $j$ ” are integers between 0 and  $n$ , inclusive, and “ $r$ ” is an identifier not previously used, then



is a network, and  $r$  is a rule node.

(Syn.8)  $r$  is the objective corresponding to the proposition that there is a relevant connection between propositions  $m_1, \dots, m_n$  such that at least  $i$  and at most  $j$  of them are simultaneously true. (Rule  $r$  is called *AND-OR* and is a unified generalization of negation ( $i = j = 0$ ), binary conjunction ( $i = j = 2$ ), binary inclusive disjunction ( $i = 1, j = 2$ ), binary exclusive disjunction ( $i = 0, j = 1$ ), etc.)

### 5.3. Fully Intensional Knowledge Representation in SNePS.

One standard way to see how a semantic theory works is to test it on various “puzzles” (cf. Russell 1905). We shall give CASSIE the following puzzle (from Castañeda, forthcoming; cf. Castañeda 1984):

The following seven statements, according to standard principles of logic, imply a contradiction:

- (1) At the time of the pestilence, Oedipus believed that: Oedipus’s father was the same as his own father but the previous King of Thebes was not the same as his own father;
- (2) Oedipus’s father was the same as the previous King of Thebes;
- (3) It was not the case that at the time of the pestilence Oedipus believed that: the previous King of Thebes was the same as his own father but the previous King of Thebes was not the same as his own father.
- (T1) For any individuals  $x$  and  $y$ : if  $x$  is (genuinely or strictly) identical with  $y$ , then whatever is true of  $x$  is true of  $y$ , and vice versa.
- (T2) The sentential matrix occurring in (1) and (3), namely: ‘at the time of the pestilence, Oedipus believed that: \_\_\_\_\_ was the same as his own father but the previous King of Thebes was not the same as his own father’, expresses something true of (a property of) the individual denoted by the singular term that by filling the blank in the matrix produces a sentence expressing a truth.
- (T3) The expression ‘was the same as’ in (2) expresses genuine or strict identity.
- (T4) The singular terms ‘the previous King of Thebes’ and ‘Oedipus’s father’ have exactly the same meaning and denotation in both direct and indirect speech.

In order to distinguish the sort of intensional knowledge representation that we advocate, it will be useful to compare our theory with those of Hector-Neri Castañeda (1972, 1975b, 1977) and Gottlob Frege (1892). We choose these, since, like ours, they are “fully intensional” in the sense that they do not use any notion of possible worlds. (There are, of course, possible-worlds analyses of Fregean theories, but Frege himself did not use them.)

Frege’s theory can resolve the contradiction by accepting (T1)-(T3) and rejecting (T4).<sup>6</sup> The

<sup>6</sup> (T4) is ambiguous. The intended reading is that ‘the previous King of Thebes’ has exactly the same meaning and denotation in both direct and indirect speech, and so does ‘Oedipus’s father,’ not that the two descriptions have the same meaning and denotation as each other.

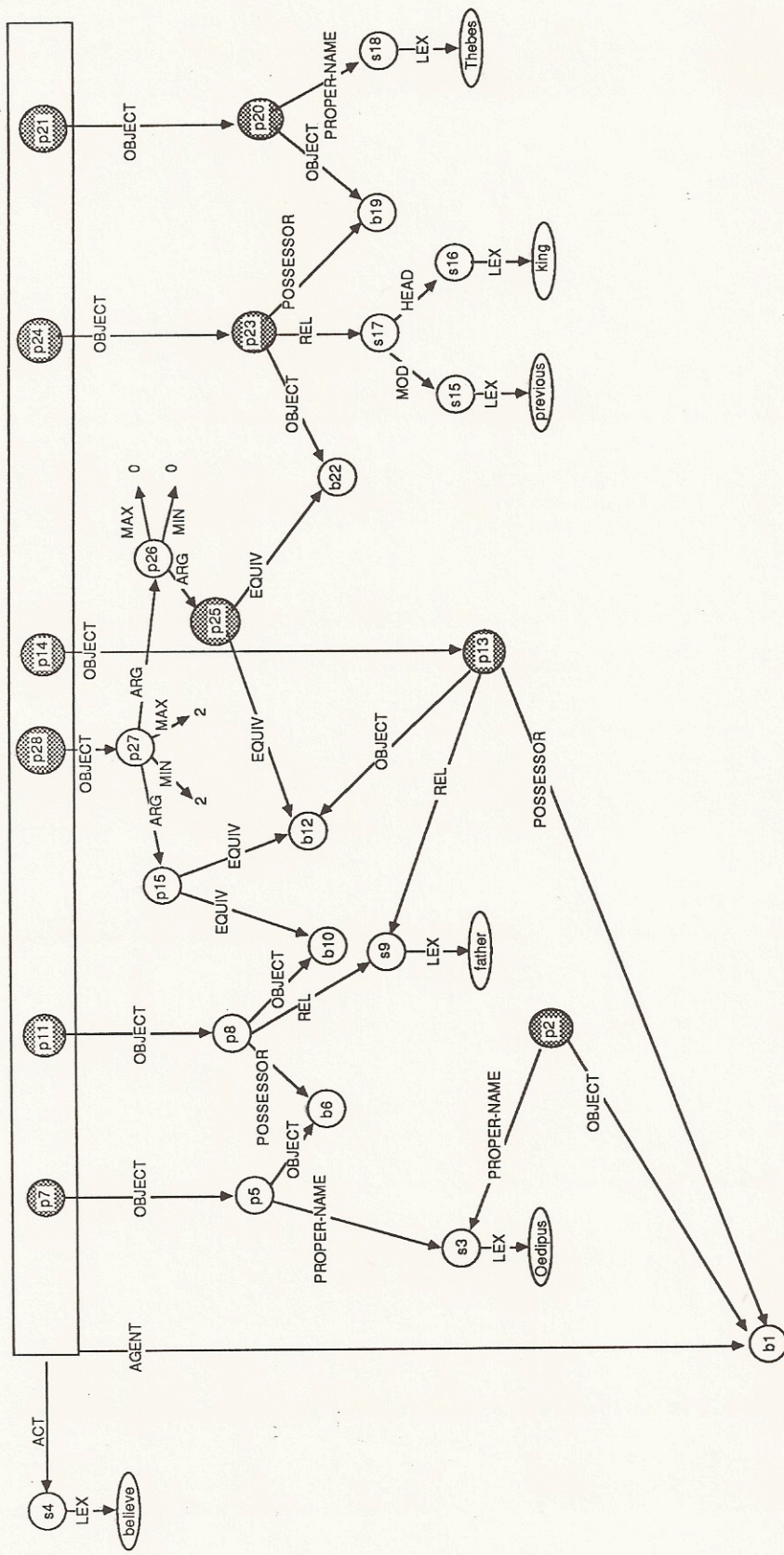


Figure 1. CASSIE's mind representing the information in sentences (1), (2), and (3). Nodes with labels beginning with "b" represent assertorally defined individuals. Nodes with labels beginning with "s" represent structurally defined individuals. Nodes with labels beginning with "p" represent propositions. Nodes with a filled-in background represent propositions that CASSIE believes. Each node in the rectangle has an AGENT arc to node b1 and an ACT arc to node s4, so the nodes with OBJECT arcs pointing to them from those nodes represent propositions that CASSIE believes Oedipus believes. The syntax and semantics of most of the representational constructs used here are defined in Shapiro and Rapaport 1986, 1987.

rejection of (T4) is accomplished by his introduction of the sense-reference distinction: On his theory, the singular terms mentioned in (T4) differ in sense and reference in direct and indirect speech.

Castañeda's theory resolves the contradiction by accepting (T1), (T2), and (T4), but rejecting (T3). The meaning of a term is always a "guise" (a kind of intensional entity that is "located" in the actual world); thus (T4) is maintained. But 'is the same as' is taken as ambiguous among several different kinds of "sameness" relations, chief of which are "consubstantiation" (*very roughly*, co-extensionality) and "consociation" (*very roughly*, co-extensionality within an intensional context). In particular, (2) is taken as expressing the consubstantiation of two guises. (For details, see Castañeda 1972, 1975b, 1977, and, especially, 1975a; cf. Rapaport 1978, 1985.)

On our SNePS/CASSIE theory, however, all of (T1)-(T4) can be accepted without contradiction. CASSIE interacts with other cognitive agents in (fragments of) natural language, interpreting each sentence in light of her previous beliefs (cf. Shapiro & Rapaport 1986, 1987; Rapaport 1986c. In this regard, our theory is similar to Hans Kamp's Discourse Representation Theory (Kamp 1984 and forthcoming; Asher 1986, 1987)). Imagine CASSIE being told sentences (1)-(3), in that order. There are several ways she can interpret the descriptions in these sentences depending on her previous beliefs. Figure 1 shows one version of a SNePS semantic network of concepts and propositions representing CASSIE's beliefs after hearing (3).

After understanding (3), CASSIE believes the propositions represented by nodes p2 (that someone—namely, the individual represented by node b1—is named 'Oedipus'), p7 (that Oedipus believes that someone—namely, the individual represented by node b6—is named 'Oedipus'), p11 (that Oedipus believes that someone—namely, the individual represented by node b10—is the father of the individual he believes to be Oedipus), p13 (that the individual represented by node b12 is Oedipus's father), p14 (that Oedipus believes that the individual represented by b12 is his father)—i.e., CASSIE believes that Oedipus believes that Oedipus's father is the same as his own father—p20 (that something—namely, the individual represented by node b19—is named 'Thebes'), p21 (that Oedipus believes that b19 is named 'Thebes'), p23 (that something, b22, is Thebes's previous king), p24 (that Oedipus believes that b22 is Thebes's previous king), p25 (that b12 = b22, i.e., that Oedipus's father is Thebes's previous king), and p28 (that Oedipus believes that p27, i.e., that b10=b12 (p15; i.e., that Oedipus's father is his own father) and that b12 ≠ b22 (p26; i.e., that his own father ≠ Thebes's previous king)—i.e., CASSIE believes that Oedipus believes that Thebes's previous king is not the same as his own father. (We are ignoring temporal information for simplicity, but without loss of generality; the representation of temporal information in SNePS is discussed in Almeida & Shapiro 1983, Almeida 1987.)

Note that there is no contradiction in this figure. To see why, let's consider each of (T1)-(T4):

(T1). In SNePS, there are two ways to represent "sameness": (a) by following the Uniqueness Principle and using a single node, or (b) by using two nodes and asserting an "EQUIV"-proposition about them (such as proposition nodes p15 and p25 of Figure 1). Only the former is genuine or strict identity; the latter is co-extensionality within a belief space (corresponding, roughly, to Castañeda's "consociation"). In SNePS, A is "true of" B if A is asserted of B. But if *x* and *y* are genuinely or strictly identical, i.e., are the same node, then surely whatever is true of one is true of the other. So SNePS/CASSIE accepts (T1).

(T2). The hidden agenda behind (T2) concerns referential opacity and transparency. Thesis (T2) asserts that the blank in the quoted context is in a referentially transparent position. Now, SNePS/CASSIE distinguishes between *de re* and *de dicto* belief reports (Rapaport & Shapiro 1984, Rapaport 1986, Wiebe & Rapaport 1986). So there are really two versions of (T2) for us to consider. On the *de re* version, the issue concerns the network fragment that consists of (a) the representation of a *de re* report of the appropriate form minus (b) the node representing whatever fills the blank: Is this fragment asserted of whatever is represented by that node? In SNePS/CASSIE, it is; so SNePS/CASSIE accepts the *de re* version of (T2). On the *de dicto* version, the issue concerns the network fragment that consists of (a) the representation of a *de dicto* report of the appropriate form minus (b) the node

representing whatever fills the blank: Is this fragment asserted of that blank filler? But we analyze *de dicto* belief reports in terms of two *de re* belief reports. So, the answer here is the same as before: SNePS/CASSIE accepts the *de dicto* version of (T2), also. Figure 1 shows the *de dicto* interpretation of (1) with respect to "Oedipus's father," so the node representing the individual denoted by the term filling the blank is b10.

(T3). Since CASSIE hears (2) *after* hearing (1), *she already has separate concepts of Oedipus's father* (represented by node b12 of Figure 1) *and the previous King of Thebes* (represented by node b22). So her representation of (2) has to be a proposition asserting that those nodes are EQUIV (represented by node p25). Hence, in this case, SNePS/CASSIE rejects (T3). However, if CASSIE had heard (2) *before* (1), her interpretation of (2) *could* be (and there is psychological evidence that it *should* be: see Anderson 1977, 1978; cf. Maida & Shapiro 1982) *two* propositions asserting of a *single* entity both that *it* was Oedipus's father and that *it* was the previous King of Thebes. In that case, SNePS/CASSIE would accept (T3).

(T4). We now come to the heart of the matter. *All nodes in the semantic network representing CASSIE's mind represent concepts in CASSIE's mind only. They never represent concepts in any other cognitive agent's mind. Concepts that are represented as being in another cognitive agent's belief space are just CASSIE's concepts about which she has beliefs that the other cognitive agent has beliefs about them.* (We have made this methodologically solipsistic point in Wiebe & Rapaport 1986. Wilks (1986: 267) makes a similar point, calling this "recursive cognitive solipsism".) Thus, (T4) is too simplistically stated to be either accepted or rejected. If the understander takes the believer to have the same concept of the referent of the singular term as the understander herself, she will use exactly the same node for this occurrence in indirect speech as she uses for occurrences in direct speech. (Figure 1 shows CASSIE making this interpretation for both "Thebes" and "the previous King of Thebes.") However, if the understander chooses to allow for the believer's having a different concept of the referent of the singular term, she will use a different node, and this occurrence in indirect speech will not have exactly the same meaning as an occurrence of the term in direct speech. (Figure 1 shows CASSIE making this interpretation for the second occurrence of "Oedipus" and for the occurrence of "Oedipus's father" in (1).)

Thus, we do not, and do not need to, reject any of (T1)-(T4) (though we do reinterpret some of them).

## 6. TOPICS IN KR FOR NLP

In this section, we present short discussions of a series of topics we believe to be important for KR for NLP. We discuss these topics in terms of SNePS/CASSIE, not because we believe that they are only relevant to her, but because we feel that we can express ourselves most clearly using her as a model.

### 6.1. Belief Spaces.

We define a believer's *belief space* to be the set of nodes (and, by extension, the concepts represented by them) dominated by proposition nodes representing propositions believed by the believer. All the nodes in Figure 1 are in CASSIE's belief space. All the nodes in Figure 1 dominated by (and including) nodes p5, p8, p13, p20, p23, and p27 are in the belief space of CASSIE's concept of Oedipus, referred to as (Oedipus CASSIE) in the notation of Rapaport & Shapiro 1984, Rapaport 1986, and Wiebe & Rapaport 1986. We say "CASSIE's concept of Oedipus" rather than just "Oedipus", because Oedipus's belief space is in his own mind, and all nodes in Figure 1 are in CASSIE's mind. (Oedipus CASSIE)'s belief space contains the nodes representing concepts that CASSIE believes to be in Oedipus's belief space. That is the sense in which (Oedipus CASSIE)'s belief space is CASSIE's model of Oedipus's belief space.

All nodes in (Oedipus CASSIE)'s belief space are also in CASSIE's belief space. *Some* of the nodes in (Oedipus CASSIE)'s belief space are such that CASSIE and (Oedipus CASSIE) have nearly the same beliefs about them. That is the case when CASSIE believes that Oedipus has a mental token representing an intensional object in *Aussersein* that closely resembles the intensional object that the node

represents for CASSIE. For example, both CASSIE and (Oedipus CASSIE) believe the proposition represented by node p20 in Figure 1. Thus, CASSIE believes that Oedipus has nearly the same concept of Thebes (node b19) as she does. Similarly, CASSIE believes that she and Oedipus closely agree on the concept of the previous King of Thebes (node b22), as well as the concept of the father relation (node s9), etc. The reason that this is close agreement rather than exact sharing of concepts is that, for example, CASSIE's concept of Thebes includes certain beliefs about its previous king that are not shared by (Oedipus CASSIE).

When CASSIE allows for the possibility that Oedipus and she might have very different concepts of an individual satisfying some description, she uses two nodes, one just in her belief space and one in (Oedipus CASSIE)'s belief space. For example, in Figure 1, node b1 represents Oedipus for CASSIE, whereas b6 represents Oedipus for (Oedipus CASSIE). Only in this way can the sentence "Oedipus believed that Oedipus's father was the same as his own father" say anything non-tautologous. "His own father" refers to (Oedipus CASSIE)'s father, represented by b12, whereas this occurrence of "Oedipus's father" refers to (Oedipus Oedipus CASSIE)'s father, represented by b10. (Oedipus CASSIE)'s belief that they are the same is represented by node p15.

It might be noticed that Figure 1 does not show p15 as being believed by (Oedipus CASSIE). Instead, (Oedipus CASSIE) believes p27, which is the conjunction of p15 and p26. Reasoning within a belief space may be carried out by the SNePS Belief Revision system (SNeBR; Martins & Shapiro 1983, 1986), assuming that all believers accept SNeBR's rules of inference. A SNeBR belief space, consisting of a set of hypotheses and all propositions derived from them, may be initialized with all propositions believed by (Oedipus CASSIE) and all propositions believed by CASSIE to be "common knowledge." SNeBR's conclusions, which in this case would include p15, could then be installed into (Oedipus CASSIE)'s belief space by making them the objects of believings by (Oedipus CASSIE). This is another way in which CASSIE can use (Oedipus CASSIE)'s belief space as a model of Oedipus's belief space.

## 6.2. Concepts of Concepts: CASSIE as Cognitive Scientist

Now that we have explicated our understanding of the relationships among mental tokens, nodes, concepts, actual objects, concepts of different cognitive agents, and concepts in different belief spaces, we will briefly discuss concepts of concepts. The first thing to note is that we have, in fact, not yet discussed concepts of concepts. We have, however, discussed concepts. That is, this paper is intended to explicate our concept of concepts. CASSIE has heretofore had beliefs about people, properties, believings, propositions, etc., but she has not had any beliefs about concepts. If we started discussing cognitive science with CASSIE, we could give her such beliefs. She would then develop a concept of concepts, and might have a theory (a coherent set of beliefs) of how concepts relate to people. In this way, concepts are just another sort of mental object on the same level as intensional people, believings, and propositions. Whether concepts in *Aussersein* are exemplified by objects in the real world is another controversy, which we do not want to pursue.

It is an intriguing idea that we could give CASSIE a theory of concepts different from the one we employed in implementing her. There would be no way for her to tell that her theory was not in accord with the facts, unless she made a prediction about her own behavior that was not confirmed, because she does not have access to the data structures used to implement SNePS.

If CASSIE thought about the issues raised in this paper, she might have nodes representing her concept of Oedipus, her concept of her concept of Oedipus, her concept of OSCAR, her concept of Oedipus as she believes OSCAR thinks of him, and her concept of OSCAR's concept of Oedipus. These would be different nodes, since they represent different intensional objects.

## 6.3. Plans

Plans are mental objects that need to be represented in NLP systems. Plans can be discussed, reasoned about, formulated, and followed.

It may be thought that any AI system that can represent and use rules, and that can evaluate a predicate by computation, rather than by inference, can just use rules to represent plans. For example,



Prolog is such a system, and one can view the Prolog rule,

$$p :- q, r, s.$$

either as the rule, “ $p$  is true if  $q$ ,  $r$ , and  $s$  are true”, or as the plan, “If you want to do  $p$ , first do  $q$ , then do  $r$ , and then do  $s$ ”. SNePS is also such a system, but in SNePS it is quite clear that rules can only be used as plans for reasoning, not as plans for acting.

To understand the problem, consider the Prolog rule,

$$p :- q, r, q.$$

This is bizarre as a rule—once  $q$  is determined as true, why bother to do so again? However, as a plan for acting, it is reasonable—there are many plans that require performing the same act more than once (for an appropriate meaning of “the same” applied to acts). The strict left-right depth-first control structure of most implementations of Prolog guarantees that the subgoal  $q$  is activated twice regardless of whether the rule is supposed to be a reasoning rule or a plan for acting. We can see that the use of rules as acting plans in Prolog depends on a pun—‘ $\vee$ ’ is simultaneously interpreted as the logical operator *AND* and as the programming-language *sequence* operator.

SNePS does not make this pun. The corresponding SNePS rule is:

$$q \vee \Rightarrow [r \vee \Rightarrow [q \vee \Rightarrow p]].$$

Although this specifies the same order of subgoal triggering as the Prolog version, SNIP (the SNePS Inference Package) would not attempt to work on  $q$  the second time. It would just reuse the results of the first effort. Therefore, SNePS rules cannot be used as plans for acting.

Is this a deficit in SNePS? No, for two reasons. First, this refusal to work on the same subgoal more than once is not only efficient, it is one of the mechanisms that allow SNIP to work with unrestricted recursive rules (Shapiro & McKay 1980; McKay & Shapiro 1981), a facility Prolog does not have. Second, the Prolog pun amounts to a semantic confusion between the logical notion of truth (or belief) and the practical notion of action.

People can explain plans to each other using their natural language. For example, I might tell a student, “To sum some numbers, choose an accumulator variable, set it to zero, and then successively add the numbers to it.” This sentence expresses a plan using certain grammatical structures to express the structure of the plan. One can then test the student’s understanding of the plan by asking questions about it. For example, one might ask, “What should the accumulator variable be initialized to when summing a sequence of numbers?”

Of course, an excellent way of demonstrating an understanding of a plan is to follow the plan in appropriate circumstances. For an NLP system, the major action carried out is the generation of language, so one might expect an NLP system to be able to understand plans about how to express itself in NL, reason about such plans, discuss them, and use them. The AI literature on planning is vast (see Vere 1987), but we know of no current NLP system that does all this.

#### 6.4. Truth.

We have persistently avoided providing a standard Tarskian denotational semantics for our networks (although nothing prevents us from doing so, in the manner, e.g., that Kamp 1984 does for his theory). For one reason, this is because the entire network is definitional (like a NIKL network; see above). For another, we are more interested with *CASSIE’s representation* of the external world than with either the external world itself or the *correctness* of her representation of it.

Again, consider Figure 1. Node p20, by its syntactic position in the network, represents a proposition that there is something named “Thebes.” The something has the properties represented by the rest of the network connected to it, and subtly changes as the network grows (cf. the “pegs” of Landman 1986). One property is that the individual represented by b22 was its previous king. Again, this may be true, false, or just confused. It may be that the previous king was in some ways alike, and in some

ways different from, CASSIE's concept of him. Would that make p23 false? Would it make p20 false?

The only construct in SNePS/CASSIE of any assertional import is the device that declares certain propositions (such as the one represented by p20) to be believed by CASSIE, and *these assertions are true by virtue of being made*—CASSIE is just built that way.

If CASSIE utters a sentence generated from a proposition she believes, we might agree or disagree with that sentence. Just as with a person, we might say the sentence is true or false, or we might say that CASSIE is confused; "Your concept of Thebes is somehow different from mine. Let's discuss it."

### 6.5. Replies to Barnden's Objections.

John Barnden (1986) makes several points that he presents as clarifications and/or objections to our intensional knowledge-representation theory as expressed in Maida & Shapiro 1982 and Rapaport & Shapiro 1984 (the reader should consult, however, Shapiro & Rapaport 1986, 1987, and Rapaport 1986 for more recent formulations). We feel that some of Barnden's points are valid complaints about lack of clarity (or timidity) in our earlier papers, some stem from his confusion about our theory, and some stem from actual differences in our theories. In this section, we will summarize Barnden's points and reply to them.

Barnden wonders whether we want our nodes to "ambassadorially represent" intensions or objects in the world, or denote one or the other (1986: 412). We are not sure we understand what Barnden means by "ambassadorially represents," but suspect that we would accept the idea that we intend nodes to ambassadorially represent intensions. We would not, however, accept Barnden's conclusion that, therefore, they *denote* objects in the world. Our ideas here are spelled out in Section 3 above.

Barnden wonders whether we want our network to model a cognitive agent, be used by a cognitive agent, or be a theoretical tool (1986: 412). The answer to this is: "all three." It is a theoretical tool for understanding cognition; it uses the methodology of Artificial Intelligence, in which one of the best theoretical tools is a computer model. In that sense, it is also a model of the mind of a cognitive agent. However, sometimes, a model of X *is* an X, and we believe that this is one of those cases (see Section 7 below). Thus, the network also is used by a cognitive agent—CASSIE.

Barnden says that belief is a relation between an extensional person and an intensional proposition (1986: 412-413) and that, therefore, "in determining what a proposition denoted by a node states, we sometimes 'dereference' the concepts denoted by argument nodes and sometimes we do not" (Barnden 1986: 413). Instead, we believe that belief is an intentional stance ascribed to people (see Dennett 1978) and therefore is a relation between an intensional person and an intensional proposition (see Section 3 above).

Barnden goes on to suppose that we would have practical problems building "a system that translates network fragments into natural-language statements. ... [W]e do not want the language generator coming out with a statement to the effect that a *concept of* Bill believes something, or to the effect that Bill believes some *truth-value*" (Barnden 1986: 413; italics in original). That our system has no such problems may be seen from the CASSIE conversation in Shapiro & Rapaport 1986, 1987. See also the discussion of natural-language expressions that refer to intensional people and propositions in Section 3 above, and the discussion of concepts of concepts in Section 6 above.

Finally, Barnden suggests that we have no consistent way of dealing with sentence sequences such as

- (1) That John is taller than Mary is Kevin's favorite proposition.
- (2) Bill believes Kevin's favorite proposition.

because the node representing the proposition would have to be dereferenced in one context and not in the other (1986: 413-414). CASSIE's belief space after hearing these two sentences is shown in Figure 2. Node p8 is built when CASSIE understands the subject clause of sentence (1). Then it is given the property of being Kevin's favorite proposition using the same constructs as were used in Figure 1 for

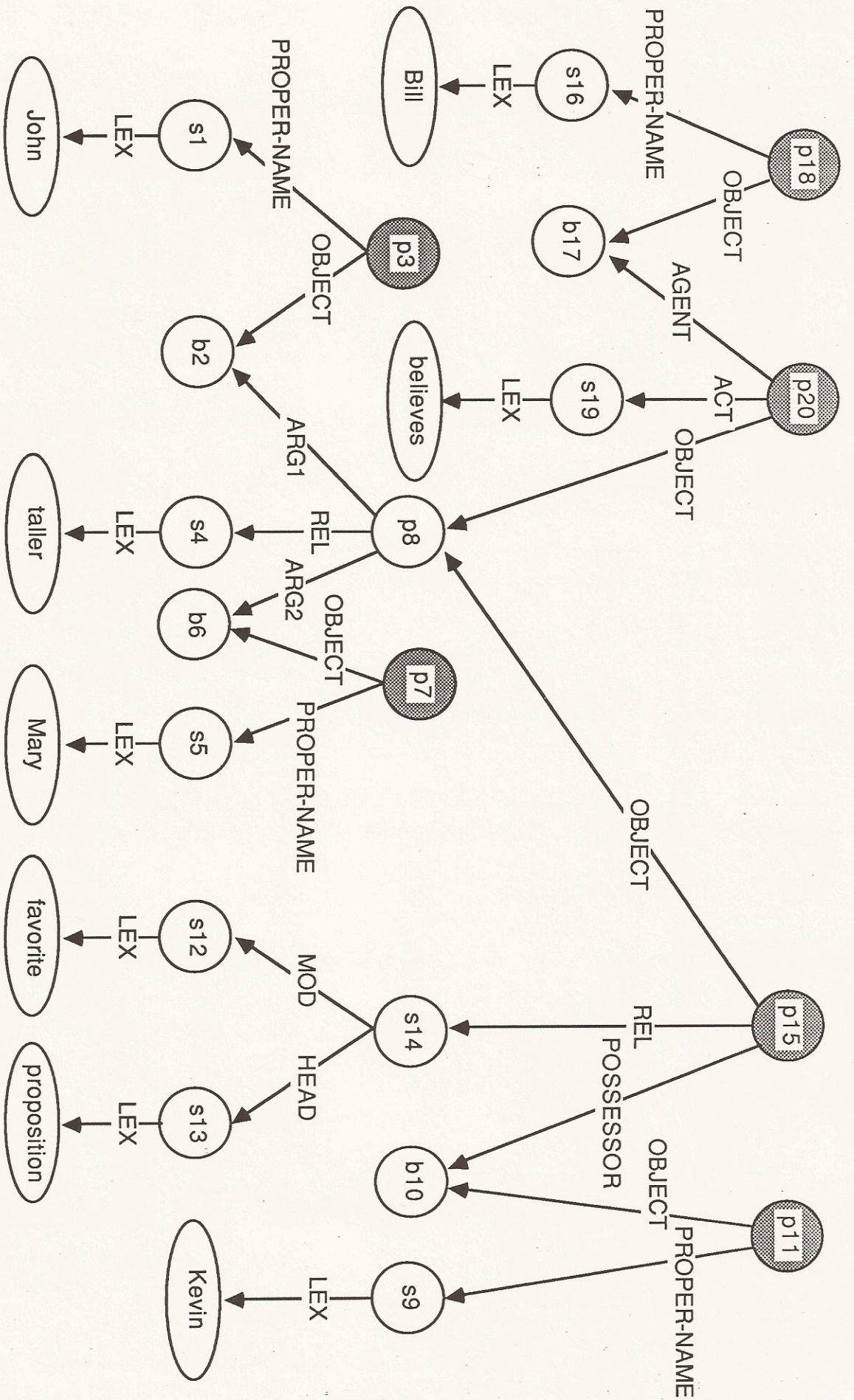


Figure 2. CASSIE's mind representing the information in the sentence sequence, "That John is taller than Mary is Kevin's favorite proposition. Bill believes Kevin's favorite proposition."

representing "the previous King of Thebes." When CASSIE hears "Kevin's favorite proposition" in sentence (2), she takes it referentially to refer to the proposition represented by p8, and attaches it as the object of Bill's believing. In any sentence in which p8 is to be mentioned nominally, CASSIE could use either the expression, "that John is taller than Mary", or the expression "Kevin's favorite proposition". (See the discussion of Lincoln's favorite proposition in Section 3 above.) Thus, two sentences that express the proposition represented by p20 are:

Bill believes that John is taller than Mary.  
Bill believes Kevin's favorite proposition.

## 7. CONCLUSIONS: BUILDING MINDS.

We have discussed knowledge representation—more accurately, we feel, the representation of beliefs—in the context of AI natural-language processing systems. KR for NLP is different from KR for robot or vision systems, because the goal of an NLP system is not to make its way in the real world, but to converse with humans about all the topics that humans discuss—real, imaginary, theoretical, and impossible.

We are now comfortable in adopting what Searle 1980 refers to as the strong AI position: an Artificial Intelligence *is* a cognitive agent and *has* a mind. Perhaps, a computer simulation of a hurricane is not a hurricane, and certainly a model of a horse is not a horse; but a wind tunnel simulation of air flow *is* air flow, a wind tunnel model of an airfoil *is* an airfoil, a wave-tank model of a wave *is* a wave, a model of a statue of a horse *is* a statue of a horse, and a computer model of a mind (if successful) *is* a mind (cf. Rapaport 1986b). As Haugeland says, "Perhaps Artificial Intelligence should be called 'Synthetic Intelligence' to accord better with commercial parlance. Thus artificial diamonds are fake imitations, whereas synthetic diamonds are genuine diamonds, only manufactured instead of dug up. . . . Despite the name, AI clearly aims at genuine intelligence, not a fake imitation" (Haugeland 1985: 255).

The task facing researchers in KR for NLP is to design a mind with the ability to handle the range and complexity of thoughts that humans have.

## REFERENCES

- (1) Almeida, Michael J. (1987), "Reasoning about the Temporal Structure of Narratives," Technical Report 87-10 (Buffalo: SUNY Buffalo Dept. of Computer Science, 1987).
- (2) Almeida, Michael J. & Shapiro, Stuart C. (1983), "Reasoning about the Temporal Structure of Narrative Texts," *Proceedings 5th Annual Conference of the Cognitive Science Society (University of Rochester)* (Hillsdale, NJ: Lawrence Erlbaum).
- (3) Anderson, J. R. (1977), "Memory for Information about Individuals," *Memory and Cognition* 5: 430-22.
- (4) Anderson, J. R. (1978), "The Processing of Referring Expressions within a Semantic Network," *Theoretical Issues in Natural Language Processing-2* (New York: Association for Computing Machinery).
- (5) Asher, Nicholas (1986), "Belief in Discourse Representation Theory," *Journal of Philosophical Logic* 15: 127-89.
- (6) Asher, Nicholas (1987), "A Typology for Attitude Verbs and Their Anaphoric Properties," *Linguistics and Philosophy* 10: 125-197.
- (7) Barr, Avron, & Davidson, James (eds.) (1981), "Representation of Knowledge," in A. Barr & E. Feigenbaum (eds.), *The Handbook of Artificial Intelligence* (Los Altos, CA: Walter Kaufmann): 141-222.
- (8) Barnden, John A. (1986), "A Viewpoint Distinction in the Representation of Propositional Attitudes," *Proceedings 5th National Conference on Artificial Intelligence (AAAI-86)*

- Philadelphia*) (Los Altos, CA: Morgan Kaufmann), Vol. 1, pp. 411-15.
- (9) Brachman, Ronald J., Fikes, Richard E., & Levesque, Hector J. (1983), "Krypton: A Functional Approach to Knowledge Representation," *IEEE Computer* 16: 67-73; reprinted in Brachman & Levesque 1985: 411-429.
  - (10) Brachman, Ronald J., & Levesque, Hector J. (eds.) (1985), *Readings in Knowledge Representation* (Los Altos, CA: Morgan Kaufmann).
  - (11) Brachman, Ronald J., & Schmolze, James G. (1985), "An Overview of the KL-ONE Knowledge Representation System," *Cognitive Science* 9: 171-216.
  - (12) Brachman, Ronald J., & Smith, Brian C. (1980), "Special Issue on Knowledge Representation," *SIGART Newsletter*, No. 70.
  - (13) Castañeda, Hector-Neri (1972), "Thinking and the Structure of the World," *Philosophia* 4 (1974) 3-40; reprinted in 1975 in *Critica* 6 (1972) 43-86.
  - (14) Castañeda, Hector-Neri (1975a), "Identity and Sameness," *Philosophia* 5: 121-50.
  - (15) Castañeda, Hector-Neri (1975b), *Thinking and Doing: The Philosophical Foundations of Institutions* (Dordrecht: D. Reidel).
  - (16) Castañeda, Hector-Neri (1977), "Perception, Belief, and the Structure of Physical Objects and Consciousness," *Synthese* 35: 285-351.
  - (17) Castañeda, Hector-Neri (1984), "Philosophical Refutations," in J. H. Fetzer (ed.), *Principles of Philosophical Reasoning* (Totowa, NJ: Rowman & Allenheld): 227-58.
  - (18) Castañeda, Hector-Neri (forthcoming), "Philosophy as a Science and as a Worldview" (Bloomington: Indiana University Dept. of Philosophy).
  - (19) Dennett, Daniel C. (1978), "Intentional Systems," in D. C. Dennett, *Brainstorms* (Montgomery, VT: Bradford Books): 3-22.
  - (20) Dennett, Daniel C. (1982), "Beyond Belief," in A. Woodfield (ed.), *Thought and Object* (Oxford: Clarendon Press): xvi-95.
  - (21) Fine, Kit (1983), "A Defence of Arbitrary Objects," *Proceedings of the Aristotelian Society*, Supp. Vol. 58: 55-77.
  - (22) Fodor, Jerry A. (1975), *The Language of Thought* (New York: Thomas Y. Crowell Co.).
  - (23) Frege, Gottlob (1892) "On Sense and Reference," trans. M. Black, in P. Geach & M. Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege* (Oxford: Basil Blackwell, 1970): 56-78.
  - (24) Hardt, Shoshana L. (1987), "Conceptual Dependency," in Shapiro 1987: 194-199.
  - (25) Haugeland, John (1985), *Artificial Intelligence: The Very Idea* (Cambridge, MA: MIT Press).
  - (26) Hobbs, Jerry R. (1985), "Ontological Promiscuity," *Proceedings 23rd Annual Meeting of the Association for Computational Linguistics (University of Chicago)* (Morristown, NJ: Association for Computational Linguistics): 61-69.
  - (27) Kamp, Hans (1984), "A Theory of Truth and Semantic Representation," in J. Groenendijk, T. M. V. Janssen, & M. Stokhof (eds.), *Truth, Interpretation and Information* (Dordrecht: Foris): 1-41.
  - (28) Kamp, Hans (forthcoming), *Situations in Discourse*, manuscript, Center for Cognitive Science (Austin, TX: University of Texas).
  - (29) King, Margaret, & Rosner, Michael (eds.) (1986), "Special Issue on Knowledge Representation," *Proceedings of the IEEE* 74: 1299-1454.
  - (30) Kramer, Bryan M., & Mylopoulos, John (1987), "Knowledge Representation," in Shapiro 1987: 882-890.

- (31) Landman, Fred (1986), "Pegs and Alecs," in J. Y. Halpern (ed.), *Theoretical Aspects of Reasoning About Knowledge* (Los Altos, CA: Morgan Kaufmann): 45-61.
- (32) Levesque, Hector J. (1986), "Knowledge Representation and Reasoning," *Annual Review of Computer Science* 1: 255-287.
- (33) MacGregor, Robert & Bates, Raymond (1987), "The Loom Knowledge Representation Language," *Proceedings of the Knowledge-Based Systems Workshop (St. Louis, MO)*; also, Technical Report ISI/RS-87-188 (Marina del Rey, CA: USC/Information Sciences Institute).
- (34) Maida, Anthony S., & Shapiro, Stuart C. (1982), "Intensional Concepts in Propositional Semantic Networks," *Cognitive Science* 6: 291-330; reprinted in Brachman and Levesque 1985: 170-189.
- (35) Martins, João. P., & Shapiro, Stuart. C. (1983), "Reasoning in Multiple Belief Spaces," *Proceedings 8th International Joint Conference on Artificial Intelligence (IJCAI-83; Karlsruhe, W. Germany)* (Los Altos, CA: Morgan Kaufmann), Vol. 1, pp. 370-373.
- (36) Martins, João. P., & Shapiro, Stuart. C. (1986), "Theoretical Foundations for Belief Revision," in J. Y. Halpern (ed.), *Theoretical Aspects of Reasoning About Knowledge* (Los Altos, CA: Morgan Kaufmann): 383-398.
- (37) McCalla, Gordon, & Cercone, Nick (1987), (eds.), *The Knowledge Frontier: Essays in the Representation of Knowledge* (New York: Springer-Verlag).
- (38) McKay, Donald P., & Shapiro, Stuart C. (1981), "Using Active Connection Graphs for Reasoning with Recursive Rules," *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI-81)*: 368-74.
- (39) Meinong, Alexius (1904), "Über Gegenstandstheorie," in R. Haller (ed.), *Alexius Meinong Gesamtausgabe*, Vol. II (Graz, Austria: Akademische Druck- u. Verlagsanstalt, 1971): 481-535; English translation ("The Theory of Objects") by I. Levi et al., in R. M. Chisholm (ed.), *Realism and the Background of Phenomenology* (New York: Free Press, 1960): 76-117.
- (40) Moser, M. G. (1983), "An Overview of NIKL, The New Implementation of KL-ONE," in C. Sidner, M. Bates, R. Bobrow, B. Goodman, A. Haas, R. Ingria, D. Israel, D. McAllester, M. Moser, J. Schmolze, & M. Vilain (eds.), *Research in Knowledge Representation for Natural Language Understanding, Annual Report*, BBN Report No. 5421 (Cambridge, MA: Bolt Beranek and Newman).
- (41) Mylopoulos, John, & Levesque, Hector J. (1984), "An Overview of Knowledge Representation," in M. L. Brodie, J. Mylopoulos, & J. W. Schmidt (eds.), *On Conceptual Modeling* (New York: Springer-Verlag): 3-18.
- (42) Rapaport, William J. (1978), "Meinongian Theories and a Russellian Paradox," *Noûs* 12: 153-80; errata, *Noûs* 13 (1979) 125.
- (43) Rapaport, William J. (1981), "How to Make the World Fit Our Language: An Essay in Meinongian Semantics," *Grazer Philosophische Studien* 14: 1-21.
- (44) Rapaport, William J. (1985), "Meinongian Semantics for Propositional Semantic Networks," *Proceedings 23rd Annual Meeting of the Association for Computational Linguistics (University of Chicago)* (Morristown, NJ: Association for Computational Linguistics): 43-48.
- (45) Rapaport, William J. (1985/1986), "Non-Existent Objects and Epistemological Ontology," *Grazer Philosophische Studien* 25/26: 61-95.
- (46) Rapaport, William J. (1986a), "Logical Foundations for Belief Representation," *Cognitive Science* 10: 371-422.
- (47) Rapaport, William J. (1986b), "Searle's Experiments with Thought," *Philosophy of Science* 53: 271-79.

- (48) Rapaport, William J. (1986c), "Syntactic Semantics: Foundations of Computational Natural-Language Understanding," in J. Fetzer (ed.) *Aspects of Artificial Intelligence* (Dordrecht, Holland: D. Reidel, forthcoming); also, *Technical Report 86-24* (Buffalo: SUNY Buffalo Dept. of Computer Science, December 1986).
- (49) Rapaport, William J. (1987), "Logic, Predicate," in Shapiro 1987: 538-44.
- (50) Rapaport, William J., & Shapiro, Stuart C. (1984), "Quasi-Indexical Reference in Propositional Semantic Networks," *Proceedings 10th International Conference on Computational Linguistics (COLING-84; Stanford University)* (Morristown, NJ: Association for Computational Linguistics): 65-70.
- (51) Russell, Bertrand (1905), "On Denoting," reprinted in *Logic and Knowledge* (ed. R. C. Marsh) (New York: Capricorn, 1956): 39-56.
- (52) Schank, Roger C. (1975), *Conceptual Information Processing* (New York: Elsevier).
- (53) Schank, Roger C., & Rieger, Charles J. (1974), "Inference and the Computer Understanding of Natural Language," *Artificial Intelligence 5*: 373-412; reprinted in Brachman and Levesque 1985: 120-139.
- (54) Searle, John R. (1980), "Minds, Brains, and Programs," *Behavioral and Brain Sciences 3*: 417-57.
- (55) Shapiro, Stuart C. (1978), "Path-Based and Node-Based Inference in Semantic Networks," in D. Waltz (ed.), *Tinlap-2: Theoretical Issues in Natural Language Processing* (New York: ACM): 219-225.
- (56) Shapiro, Stuart C. (1979), "The SNePS semantic network processing system," in N. V. Findler (ed.), *Associative Networks: The Representation and Use of Knowledge by Computers* (New York: Academic Press): 179-203.
- (57) Shapiro, Stuart C. (ed.) (1987), *Encyclopedia of Artificial Intelligence* (New York: John Wiley & Sons).
- (58) Shapiro, Stuart C., Martins, João P., & McKay, Donald P. (1982), "Bi-Directional Inference," *Proceedings Cognitive Science Soc 4*: 90-93.
- (59) Shapiro, Stuart C., & McKay, Donald P. (1980), "Inference with Recursive Rules," *Proceedings AAAI 1*: 151-53.
- (60) Shapiro, Stuart C., & Rapaport, William J. (1986), "SNePS Considered as a Fully Intensional Propositional Semantic Network," *Proceedings 5th National Conference on Artificial Intelligence (AAAI-86; Philadelphia)* (Los Altos, CA: Morgan Kaufmann), Vol. 1, pp. 278-83.
- (61) Shapiro, Stuart C., & Rapaport, William J. (1987), "SNePS Considered as a Fully Intensional Propositional Semantic Network," in McCalla and Cercone 1987: 262-315.
- (62) Smith, Brian C. (1982), "Reflection and Semantics in a Procedural Language," Technical Report MIT/LCS/TR-272 (Cambridge, MA: MIT Laboratory for Computer Science); page reference is to the prologue, reprinted in Brachman & Levesque 1985: 32-39.
- (63) Srihari, Rohini K. (1981), "Combining Path-Based and Node-Based Inference in SNePS," *Tech. Report No. 183* (Buffalo: SUNY Buffalo Dept. of Computer Science).
- (64) Stich, Stephen (1983), *From Folk Psychology to Cognitive Science: The Case Against Belief* (Cambridge, MA: MIT Press).
- (65) Vere, Steven A. (1987), "Planning," in Shapiro 1987: 748-758.
- (66) Vilain, Marc (1985), "The Restricted Language Architecture of a Hybrid Representation System," *Proceedings Ninth International Joint Conference on Artificial Intelligence (IJCAI-85; Los Angeles)* (Los Altos, CA: Morgan Kaufmann): 547-551.
- (67) Weischedel, Ralph M. (1986), "Knowledge Representation and Natural Language Processing," *Proceedings of the IEEE 74*: 905-920.

- (68) Wiebe, Janyce M., & Rapaport, William J. (1986), "Representing *De Re* and *De Dicto* Belief Reports in Discourse and Narrative," *Proceedings of the IEEE* 74: 1405-1413.
- (69) Wilks, Yorick (1986), "Relevance and Beliefs," in T. Myers, K. Brown, & B. McGonigle (eds.), *Reasoning and Discourse Processes* (London: Academic Press): 266-289; earlier version pre-printed as "Relevance, Points of View and Speech Acts: An Artificial Intelligence View," Memoranda in Computer and Cognitive Science No. MCCA-85-25 (Las Cruces: New Mexico State University Computing Research Laboratory).
- (70) Winograd, Terry (1975), "Frame Representations and the Declarative/Procedural Controversy," in D. G. Bobrow & A. M. Collins (eds.), *Representation and Understanding: Studies in Cognitive Science* (New York: Academic Press): 185-210; reprinted in Brachman & Levesque 1985: 358-370.