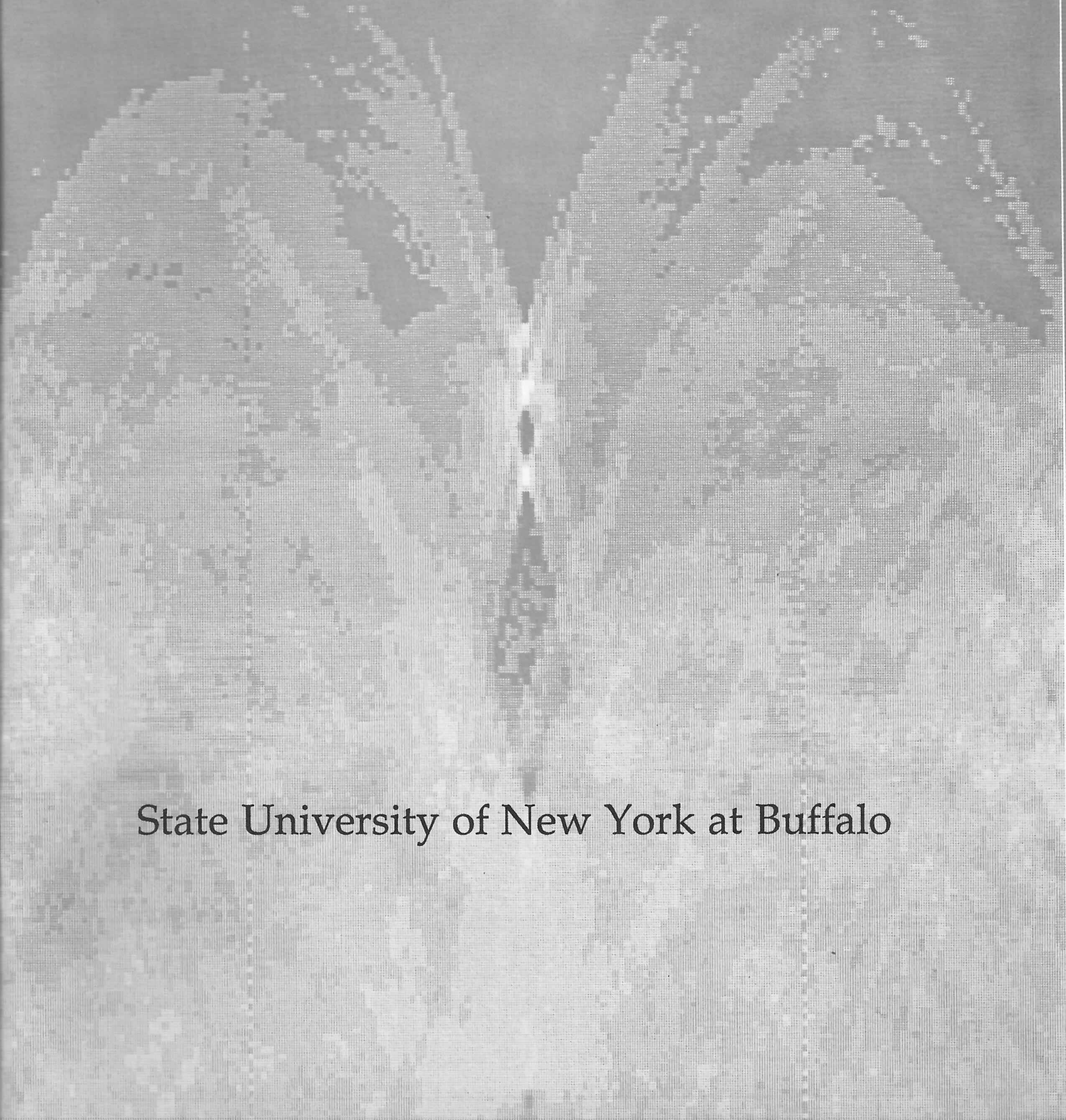
COMPUTER SCIENCE

RESEARCH REVIEW 1987—88



State University of New York at Buffalo

Intensional Knowledge Representation

William J. Rapaport, Janyce M. Wiebe, and Randall R. Dipert

Abstract

This article discusses intensional knowledge-representation and reasoning as a foundation for modeling, understanding, and expressing the cognitive attitudes of intelligent agents. In particular, we are investigating both "representational" and "pragmatic" issues: The representational issues include (1) the design of representations rich enough to support the interpretation and generation of referring expressions in opaque (i.e., intensional) contexts, to be accomplished by means of structured individuals and the notion of "belief spaces", and (2) the design of representations rich enough to support the use of intentions and practitions for representing and reasoning about action. The pragmatic issues include the recognition of a speaker's intentions (for interpreting referring expressions in opaque contexts) and the generation of referring expressions in opaque contexts based on the intentions of the cognitive agent. This pragmatic part of the overall project uses the results obtained from our representational work on intentions and practitions. The research is of significance for natural-language processing and computational models of cognition and action.

1. Background.

We are investigating intensional knowledge-representation and reasoning as a foundation for modeling, understanding, and expressing the cognitive attitudes of intelligent agents. The representational issues include (*i*) the design of representations rich enough to support the interpretation and generation of referring expressions in opaque (i.e., intensional) contexts, to be accomplished by means of structured individuals and the notion of "belief spaces", and (*ii*) the design of representations rich enough to support the use of intentions and practitions for representing and reasoning about action. The pragmatic issues include the recognition of a speaker's intentions (for interpreting referring expressions in opaque contexts) and the generation of referring expressions in opaque contexts based on the intentions of the cognitive agent. This pragmatic part of the overall project will use the results obtained from our representational work on intentions and practitions.

In our current NSF-funded project, "Logical Foundations for Belief Representation", we are designing and implementing a logically and psychologically adequate computer system capable of representing and reasoning about the cognitive attitudes of intelligent agents, together with a firm theoretical foundation for the representation in terms of a fully intensional semantics. The agents include users, other AI systems, and the system itself; the cognitive states include beliefs, knowledge, goals, intentions, and desires. The system is able to represent nested beliefs, and is sensitive to the intensionality and indexicality of beliefs, in particular, to the phenomenon of "quasi-indexicality", a feature at the core of self-referential beliefs (see below). Finally, the system is able to expand and refine its beliefs by interacting with users in conversational situations. The system is being implemented in the SNePS propositional Semantic Network Processing System [44,46,47] using an Augmented Transition Network (ATN) grammar for interpretation and generation [45].

SNePS is a *propositional* (as opposed to object-oriented) semantic-network knowledge-representation and reasoning system, in which the following conditions hold:

(a) each node represents a unique concept;

(b) each concept represented in the network is represented by a node;

(c) Uniqueness Principle: each concept represented in the network is represented by a unique node;

(d) arcs represent non-conceptual, binary, structural relations between nodes;

(e) the knowledge 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 (cf. [25,34,46,47]). Further, if we view SNePS as a system for modeling the mind of a cognitive agent, then all represented concepts are in the mind of the cognitive agent. For ease of exposition, we have named this cognitive agent 'CASSIE' (for *C*ognitive *A*gent of the *S*NePS *S*ystem—an *I*ntelligent *E*ntity).

A central notion in our system is that of "belief space". A cognitive agent's belief space consists of the propositions the agent believes and the concepts that make up these propositions. For example, if CASSIE believes that John believes that Lucy is rich, then that Lucy is rich is a belief that appears in CASSIE's model of John's belief space. CASSIE's own belief space consists of the propositions that are currently "asserted" and the concepts that make up these propositions. (Cf. [32,37,48]. Belief spaces of varying kinds are also used by [23,24,49,50] among others. See [39] for a survey.) One major difference between our approach to the theory of belief spaces and those of some others is that we represent the (contents of the) mind of a *single* agent, namely, CASSIE. Therefore, any belief space that is represented in addition to CASSIE's contains those beliefs that *CASSIE* believes are held by the other agent; it does not represent the other agent's beliefs directly.

We are extending the intensional knowledge-representation facilities of SNePS to handle such diverse cognitive-attitude expressions as 'know that', 'hope that', 'want', and 'intend to'. The first of these has required a revision of our earlier representation for quasi-indexical *de dicto/de se* belief reports in order to enable the implementation of the rule of veridicality (that knowledge implies truth); this is documented in [41] and described briefly, below.

We represent a *de re* belief report expressed in English by

(1) John believes *of* Lucy *that* she is rich

as (roughly)

(1A) John(m_1) & Lucy(m_2) & Believe(m_1, Rich(m_2)),

where the m_i can be treated as Skolem constants; they are, in fact, nodes in a SNePS semantic network (m_1 represents CASSIE's concept of an object named 'John', and m_2 represents CASSIE's concept of an object named 'Lucy'). Thus, (1A) expresses the proposition that something (viz., m_1) is named 'John', that something else (viz., m_2) is named 'Lucy', and that m_1 believes of m_2 that it (m_2) is rich.

We represent a *de dicto* belief report expressed in English by

(2) John believes *that* Lucy is poor

as (roughly)

(2A) John(m_1) & Believe(m_1, Lucy(m_2)) & Believe(m_1, Poor(m_2))

(That is, something (viz., m_1) is named 'John' and it (m_1) believes of something else (viz., m_2) that it (m_2) is named 'Lucy' and is poor.)

A quasi-indicator is an expression within an intentional context that represents a use of an indicator (i.e., an expression that makes a strictly demonstrative reference) by another person. Quasi-indicators pose problems for natural-language representation and reasoning systems, because—unlike pure indicators—they cannot be replaced by co-referential noun phrases without changing the meaning of the embedding sentence. Therefore, the referent of the quasi-indicator must be represented in such a way that no invalid co-referential claims are entailed. We formerly represented a quasi-indexical *de dicto/de se* belief report expressed in English by

(3) John believes that he* [i.e., he himself] is rich

as (roughly)

(3A) $\text{John}(m_1)$ & $\text{Believe}(m_1, \text{Ego}(m_2))$ & $\text{Believe}(m_1, \text{Rich}(m_2))$,

where 'Ego(x)' represents that x is (CASSIE's representation of) the believer's "self-concept"; it means (again, roughly) "x is me". (In each case, the actual representations used are shown in the semantic networks of Figure 1.) However, this original representation of quasi-indexical *de se* belief reports has to be modified to solve a problem first observed by Castañeda, namely, that the simple rule of veridicality, '(A knows that P) implies P', apparently does not hold if P contains a quasi-indicator. Thus, to implement veridicality, both (3) and the analogue of (3) for the cognitive attitude of knowledge must be represented as (roughly)

(3B) $\text{John}(m_1)$ & $\text{Believe}(m_1, \text{Rich}(m_1))$
(3K) $\text{John}(m_1)$ & $\text{Know}(m_1, \text{Rich}(m_1))$

(see Figure 2). Otherwise, from the analogue of (3A) for knowledge, CASSIE would infer by veridicality that *she** (i.e., she herself) was rich. Furthermore, we found a single rule that holds for all propositions P, including quasi-indexical ones, and we provided support for the necessity of considering sentences in the context of extended text (e.g., discourse or narrative) in order to fully capture certain features of their semantics. (Cf. [41].)

We have also provided a formal syntax and semantics for SNePS considered as the (modeled) mind of a cognitive agent such as CASSIE. The semantics is based on a Meinongian theory of the intensional objects of thought that is appropriate for AI considered as "computational philosophy" or "computational psychology". (See [46,47].) The Meinongian theory, providing a foundation for "epistemological" or "naive" ontology, has itself been studied in some detail (cf., most recently, [35].)

As part of an investigation into the philosophical implications of our research, we have considered what it means to understand natural language and whether a computer running an AI program designed to understand natural language (such as our SNePS program) does in fact do so. We claim that a certain kind of semantics is needed to understand natural language, that this kind of semantics is mere symbol manipulation (i.e., syntax), and that, hence, it is available to AI systems. This has also resulted in a clearer understanding of the semantics of our LEX arcs. (See [38].)

2. Disambiguation of Belief Reports.

The main unsolved problem in our current research is the disambiguation of *de re* and *de dicto* belief reports. The problem is this: We currently require the user to indicate to CASSIE whether a belief report is *de re* or *de dicto*; this is done by using one of the two canonical forms:

S believes that N is F	(for *de dicto*)
S believes of N that s/he is F	(for *de re*)

(where S names a cognitive agent, N is a proper name or definite description, and F names a property). But, of course, people do not ordinarily speak this way. Rather, they would most likely use other forms for *both* kinds of reports—e.g., "believes that" or "thinks that", etc., or even without an explicit intentional verb (as in narrative, where, e.g., the narrator reports the character's belief using direct quotation). The issue, then, is this: on what basis should CASSIE determine which kind of report was intended by the speaker? This clearly involves some beliefs on the part of CASSIE about the *speaker's* beliefs about the individuals being spoken of.

However, the distinction between *de re* and *de dicto* is not as clear-cut as many researchers suppose. In [48], we investigated the disambiguation of belief reports as they appear in discourse and narrative. In the research described in Section 1, above, the distinction between *de re* and *de dicto* belief reports was made solely on the basis of their representations. However, this analysis is sufficient only when belief reports are considered in isolation. This is because a representational structure built as the result of CASSIE's understanding, say, a *de dicto* report might be modified by a later interaction to incorporate *de re* information (or vice versa). Thus, in order to sufficiently represent *de re* and *de dicto* belief reports as they appear in discourse and narrative, more complex belief structures need to be considered. Further, we cannot meaningfully apply one, but not the other, of the concepts *de re* and *de dicto* to these more complex belief structures. We maintain that the concepts *de re* and *de dicto* do not apply to an agent's conceptual representation of her beliefs; rather, they apply to the utterance of a belief report on a specific occasion. A cognitive agent interprets a belief report such as "*S* believes that *N* is *F*", or "*S* said, '*N* is *F*'" *de dicto* if she interprets it from *N*'s perspective, and she interprets it *de re* if she interprets it from her own perspective.

Thus, the *de re/de dicto* distinction (under our analysis) is largely a pragmatic issue that pertains to the *use* of language, rather than strictly a representational issue. This can be highlighted by a different terminology that reflects this change of perspective: a *de re* report is one that contains a referring expression that is the *cognitive agent's* reference, whereas a *de dicto* report is one that contains a referring expression that is the *believer's* reference. (Cf. [9]'s notion of "speaker's reference" vs. "believer's reference".)

Based on this new theory of the distinction as it concerns natural-language *understanding*, we turn to the following problem: how are different interpretations of each type of referring expression distinguished in actual language use?

First, the speaker's intentions are crucial. When CASSIE understands a cognitive-attitude report, she should recognize the speaker's intentions concerning the perspective from which to understand the referring expressions. When CASSIE generates a cognitive-attitude report, she is the "speaker" and should herself have the intention to generate referring expressions from some particular perspective. Therefore, CASSIE should recognize the speaker's intentions while *parsing* (interpreting) cognitive-attitude reports, and CASSIE's own intentions should underlie her *generation* of cognitive-attitude reports. In order for this to happen, we need to endow CASSIE with intention, and treat generation as an action carried out to fulfill an intention. Accordingly, this part of our research will use the results obtained from the part concerned with intention and action (see Sect. 4, below).

A second issue is that CASSIE's current beliefs impose constraints on possible ways she can express (generate) a cognitive attitude of a third agent, and a third issue is that CASSIE should reason from her current beliefs to weigh the likelihood of one interpretation over another. These last two issues are probably subproblems of the first: the second is a constraint on the intention CASSIE will form; the third is part of recognizing the speaker's intentions. This, in turn, is a special case of the more general need for representing and reasoning about intentions in natural-language understanding (especially generation) and in planning (see below, Sect. 4).

3. De Re and De Dicto Interpretations of Referring Expressions.

It follows from what we said in the previous section that names and descriptions that appear within opaque contexts can be *interpreted* from the listener's perspective (i.e., from her belief space) or from the listener's view of another agent's perspective (i.e., from a nested belief space). Similarly, such names and descriptions must be *generated* from either the speaker's belief space or from a nested belief space. For example, third-person narrative probably indicates an intention that cognitive-attitude reports be interpreted from CASSIE's (i.e., the reader's or hearer's) perspective. On the other hand, direct speech (or thought) of a character indicates that that character's is the appropriate perspective from which to interpret referring expressions. We plan to analyze these phenomena in ordinary language, as well as the relevant factors mentioned in the discussion of speaker's intentions, above, in order for CASSIE to be able to make appropriate interpretations of, and appropriately generate, names and descriptions in opaque contexts. Much of the data for this part of the project will be drawn from work we are doing in conjunction with members of the SUNY Buffalo Graduate Group in Cognitive Science on "deictic centers" in narrative (cf. [8]).

3.1. Assertional and Structural Information.

In a SNePS network, arcs *pointing to* a node represent information *asserted* about the node, whereas arcs *emanating from* a node represent *structural* information about the node. E.g., the representation of 'Rover is a dog' might be a node, *m*, with a MEMBER arc pointing to a node representing Rover and a CLASS arc pointing to a node representing the class of dogs (see Figure 4a; cf. [46,47]). The *structural* information about the proposition node, *m*, is contained in the MEMBER and CLASS arcs. *Assertional* information about Rover and about the class of dogs is contained in node *m*.

The distinction between structural and assertional information in semantic networks can be found in [51]. A similar distinction has long been present in the philosophical literature under the heading "internal and external predication". Using the latter terminology, assertional information is *externally* predicated of an object, and structural information is *internally* predicated. (Cf. [10,27,29,31,33,42].)

In the network of Figure 4a, the Rover-node is simply a node labeled 'Rover'. Even if—as it in fact would be in our use of SNePS—it were a node *asserted* to be named 'Rover' (by means of an OBJECT–PROPER-NAME case frame), it would be considered an *unstructured* individual (Fig. 4b). If it had a LEX arc to a sensory node labeled 'Rover', it would be a *structured* individual (Fig. 4c; cf. [25,45-47]). Another example of a structured object would be an object that was "constituted" (in the Meinongian sense; cf. [29,34]) by the properties of being red and being square—viz., a red square. Proposition nodes are also structured objects, and a proposition node that is in a nested belief space also has assertional information about it represented in the network, namely, that some agent believes it.

3.2. The Representation of Assertional and Structural Information.

Indefinite descriptions can be interpreted to be specific, non-specific, or generic. An indefinite description that is interpreted to be *specific* is taken to refer to a particular individual, whereas an indefinite description interpreted to be *non-specific* is not taken to refer to any particular individual. For example, under the specific reading of 'John wants to catch a unicorn', there is a specific unicorn John wants to catch, perhaps the one he just read about who escaped from the circus. Under the non-specific reading, John is unicorn hunting; any old unicorn will satisfy him. An indefinite description interpreted to be *generic* refers to the class of entities, in the sense, e.g., that under the generic reading of 'a unicorn is a mythical animal', 'a unicorn' refers to the class of unicorns as a whole.

Definite descriptions can be interpreted referentially or attributively [18,19]. A definite description interpreted referentially is taken to refer to a particular entity. A definite description interpreted attributively is used to say something about whoever or whatever satisfies the description. For example, on the referential reading of 'the strongest man in the world',

> The strongest man in the world can lift 440 lbs.

says that the particular man who happens to be the strongest man in the world can lift 440 lbs. On the attributive reading, it says that whoever is the strongest man in the world can lift 440 lbs.

We are investigating a scheme under which (following [20,21]) similarly structured entities represent, on the one hand, indefinite descriptions interpreted (or used) specifically and definite descriptions interpreted (or used) referentially and, on the other hand, indefinite descriptions interpreted (or used) non-specifically and definite descriptions interpreted (or used) attributively. The referring expressions in the first category—those taken to refer to specific entities—will be represented by non-structured individuals about which the relevant descriptions are asserted; the latter—those not taken to refer to specific entities—will be represented by structured individuals that are defined by, or comprised of, the descriptions, but that carry no assertional import in themselves (cf. [48]). These representations will allow us to control substitutional inferences, and, together with our representation of belief spaces, they predict certain interactions between the *de re/de dicto* distinction (as analyzed in [35,40,48]) and non-specific and attributive interpretations of noun phrases.

3.3. Structured Individuals and Belief Spaces.

The objects of belief are propositions, which carry assertional information. Structured individuals are structured concepts and, as such, carry no assertional information (cf. [29]). Thus, they are not objects of belief. Nested belief spaces are *CASSIE*'s models of what others believe; they are not *our* models of the other agent's concepts. In a nested belief space, it is appropriate to include a description D asserted about an individual, since this represents another's belief that someone is D. However, belief spaces are not intended to model what being D (or the property of D-ness) means to other agents. Thus, although the *structure* of a structured individual may involve propositions that are nested within some agent's belief space, the structured concept itself is not a proposition included in a belief space. Consider the reading of

(4) Nadia wants a dog like Ross's,

in which D = 'a dog like Ross's' is interpreted non-specifically. We suggest that D is not itself subject to the *de re/de dicto* distinction, although the propositional information that Ross has a dog and that Ross is named 'Ross' may be: the interpretation of D is not based on a belief about some individual (cf. [48]). Rather, D is interpreted to be a structured concept of dog-like-Ross's-ness. Constrast this reading with the reading of (4) on which D is interpreted specifically (i.e., Nadia wants some particular dog). Then, D *is* itself subject to the *de re/de dicto* distinction, since its interpretation is based on a belief about some individual, which can be included in a nested belief space.

In accordance with our analysis of the *de re/de dicto* distinction presented in [32,35,37,40] we suggest that indefinite descriptions used non-specifically and definite descriptions used attributively—represented by structured individuals—are *not* subject to the *de re/de dicto* distinction. On the other hand, indefinite descriptions used specifically and definite descriptions used referentially—represented by non-structured individuals—*are* subject to it.

3.4. Acquisition and Revision.

Structured individuals are not proposed as conceptual primitives; therefore, their acquisition in the context of discourse and narrative is being investigated, as is the relationship between corresponding structured and non-structured individuals. The cognitive agent may form a structured individual of a description D and then later form a concept of a non-structured individual about which D is asserted; she may also form these concepts in the reverse order. This will affect the structure of the concepts as well as the cognitive agent's understanding and generation of the corresponding referring expressions.

4. Intentions and Practitions.

There is an extensive literature in artificial intelligence on the topic of "planning" (cf. [12]). In related research, one of us (Dipert) has proposed that deliberation and agent-reasoning in addition to planning, are required for an entity justifiably to be said to act [15-17].

It is a commonplace in philosophy, law, and everyday moral discourse that human action requires intentions at some point. With a slight danger of oversimplification, we can describe most of the extensive literature on planning as being concerned with the simulation of rational behavior—of behavior that is somehow caused by processes associated with various and complex "rational" inferential patterns. However, it seems to be a strong and long-held intuition in philosophy and various fields of normative inquiry that such a process of inferences—however rational—"causing" behavior is insufficient for genuine action—that is, to be able to speak of an agent (intentionally) acting. Current research in action theory is virtually unanimous in maintaining that genuine actions are not simply rational behavior (4,11,13)—in fact, one could argue that many intentional actions are not even rational. (On rational behavior, cf. also [5]-[7].

What seems necessary for something to constitute an action is a state, wedged between the activities of planning and overt behavior, of intending to do it. The "it" that we intend to do requires a careful, and possibly non-propositional, characterization. We suggest that intending to do something requires certain previous processes—such as deliberating, however summarily, whether to do it. Finally, it is possible that there are actions that are in no sense planned, though we intentionally perform them. These might be immediately executable intentions or intentional behavior that we have no inclination or time consciously to plan [13].

4.1. Background.

An artificial system (e.g., CASSIE) that can represent objects having all the significant structural features of (human) thoughts, that can have distinct (propositional) attitudes toward these entities, that can manipulate them in a way significantly similar to what we call "reasoning" in human cognitive agents, but that cannot intentionally act, will fail to mimic significant features of human agents:

(a) The system will not be responsible for its behavior, in the sense that is often invoked in everyday discourse (especially legal discourse). The system "behaves" in the sense in which rocks fall to the ground.

(b) The system may produce utterances, but could not be said to speak, since speech acts presumably require an intention. (Perhaps the system could not even be said to mean anything with its utterances, since this too requires intentions.)

(c) If the system lacks the ability to represent the intentions and practical reasoning of others, it cannot be said to understand the ubiquitous pragmatic elements of others' speech [22]. In fact, the system will not experience any artifacts (art works, tools, symbols), as humans frequently do (cf. [15]).

(d) As described above (Sects. 1-3), the system will have beliefs about propositions. But these beliefs may not be rational: they might not be the product of reason, since reason itself might require an intention to consider or contemplate a proposition, and an intention to infer other propositions from it. Without such intentional control over our thought processes, it is unlikely that we are reasoning; we are instead producing apparently intelligent new thoughts, somehow. Without intentions, "thought", inference, and planning would be at the level of behavior, not at the level of action.

Although (a) has generally been ignored in AI research, (b) and especially (c) have been partially addressed in [1] and, more recently, [28]; (d) is at least suggested by the expression 'meta-rule' [14]. What has been missing is a fundamental analysis of intentional action: what its essential features are for us as actors and, derivatively, what it is we attribute to other agents when we regard them as having acted. (For a discussion of "agency" within cognitive science, but which diverges from our approach at many points, see [4].)

4.2. The Importance of Acting and Intending.

It is fairly obvious that the major issues involved in the adequate representation of action and other consciously goal-directed behavior include all of the issues from the other domains of knowledge representation: the proper representation of "rules", possibly using non-standard connectives or second-order techniques; the need for representing intensional objects; and the need for representing quasi-indexicality and other objects of thought reflected in speech by the use of indexicals. All of these either have been or are being implemented in SNePS.

Additionally, it is a theme in the philosophical literature on action theory that the objects of action-directed attitudes, such as "intending," have distinctively *non-propositional* characteristics. [11] has called these objects of thought that are analogous to propositions but that are associated with activities, 'practitions'. (Cf. [26] for an AI application of practitions to legal reasoning.)

4.3. The Objects of Intending.

Toward what kind of object are the mental activities of intending and other "practical" attitudes directed? In speaking of "practical" attitudes, we refer to cognitive processes that are directed to eventual action upon the world, as opposed to the propositional attitudes involved in understanding it. The objects of intending surely have some of the features that propositions have. They are intensional, thus typically barring substitution of identicals or of logical equivalents: If I intend to sue the owner of the company that produced this leaky pen, and, unbeknownst to me, my beloved sister is the owner of the company, it does not follow that I intend to sue my sister. It is only under a description, or through what [10,11] calls a "guise" of the object, that we have a cognitive relationship to it.

But are the objects of intending exactly propositions? As Castañeda has argued extensively [11], there are reasons to think that this is not the case:

(1) The surface grammatical structure of the objects of most practical attitudes is an infinitive construction, whereas the surface grammatical structure of the object of most propositional attitudes is a sentence (the asterisks indicate ill-formed expressions):

> I intend to swim tomorrow.
> I believe that I will swim tomorrow.
> * I intend that I will swim tomorrow.
> * I believe to swim tomorrow.

(2) The "value" toward which a cognitive agent directs his or her attention in the object of a practical attitude is not a truth value. From 'I intend to swim tomorrow', nothing can be derived about what I believe to be "true" about swimming tomorrow. In particular, I need not even believe that I will swim tomorrow. The cognitive emphasis in a practical attitude, we might metaphorically say, is on the activity itself, not on the state that results from that doing, nor even on the fact of the doing. Castañeda calls this other, practical, value "legitimacy". Truth serves as an ideal for believing propositions in the way that "legitimacy" serves as an ideal for practical attitudes.

(3) Castañeda uses the term 'practition' for the non-propositional entities that serve as objects of practical attitudes, that consequently appear as infinitive constructions, and that are assessed by an agent in terms purely of "legitimacy". The often implicit object serving as subject of the infinitive construction is bound to the predicate by a practitional copula, distinct from the propositional copula. One is thus never tempted to consider whether the practition expressed by 'Sally to swim' is true or false, although one may be tempted to consider whether it is "legitimate" or not.

(4) Practitions can be made the embedded objects of propositional attitudes by the addition of certain operators. From contemplating the "legitimacy" of the *practition*:

> Sally to swim tomorrow,

we might be led to entertain the *proposition*:

> It is obligatory for Sally to swim tomorrow, since she promised,

after which we might conclude:

I believe that it is obligatory for Sally to swim tomorrow, since she promised.

In this example, we see a constituent that is crucial for deliberation—the choice among conflicting, contemplated activities. Namely, "legitimacy" is typically fragmented for the sophisticated acting agent into many species: "legitimacy" that derives from having promised to do something, from following federal law, from following state law, from strictly prudential (self-interest) grounds of numerous sorts and presumably relativized to diverse end-goals, from moral theory or moral intuitions, and so on.

(5) Logical connectives connect practitions to practitions, practitions to propositions, and, of course, propositions to propositions to make new practitions or propositions (cf. [30]). In the following examples, 'Obligatory' serves as an operator on a *practition* and results in a *proposition*, whose subscript (here left unspecified as 'k') represents the specific kind of "legitimacy" the practition or proposition is held to have. For instance, the sentence

Sam is obliged$_k$ to drive me to the airport tomorrow

expresses a *proposition* whose interpretation expressed in a canonical form is:

Obligatory$_k$ (Sam to drive me to the airport tomorrow),

where the embedded *practition* is:

Sam to drive me to the airport tomorrow.

The sentence

If it is not raining tomorrow, then Sam is obliged$_k$ to drive me to the airport tomorrow

expresses a proposition whose interpretation is:

Obligatory$_k$ (If it is not raining tomorrow, then Sam to drive me to the airport tomorrow)

and not:

If it is not raining tomorrow, then Obligatory$_k$ (Sam to drive me to the airport tomorrow).

(Extensive reasons for the former interpretation as standard are given in [11].)

4.4. Intending and Acting.

The attitude of intending has certain key functional features: it leads directly to behavior, and it must arise through certain processes involving other propositional and non-propositional attitudes.

Once a particular obligation sentence is contemplated with the thought of execution, a complex process is begun, which we term "practical reasoning". Although they do not necessarily occur sequentially, we can identify three distinct subprocesses in practical reasoning: planning, deliberation, and agent-reasoning. Planning is the process of articulating goals into smaller and more easily executable steps using means-ends reasoning. Deliberation is the process of searching for competing value beliefs and practitions and weighing conflicting such items in the case of clashes. Finally, agent-reasoning is reasoning about the subject to whom the obligatory activity is to be assigned. If the resulting practition has as its agent the reasoning agent him- or herself, the result is an *intention*. Otherwise, the result is a *prescription*—which might lead to giving a command or to other actions directed toward coercing or inducing an agent to fulfill the obligation assigned to them.

4.5. Implementation Status.

Modest sets of case frames representing the conceptual structure of an agent's propositional and practitional thoughts have been implemented in SNePS and in PSNePS (see below, Sect. 5). Both forward- and backward-chaining inferences have been performed using rules for planning (special practical versions of modus ponens), as well as elements crucial for deliberating and agent-reasoning. Both deliberative search (allowing two instances of backward chaining, but no more) and decisions involving clashes (with a predetermined partial-ordering of kinds of obligatoriness, thus giving rise to genuine dilemmas that would not occur were the ordering linear, and which are decided pseudo-randomly) have been implemented.

Currently, we are planning the following:

(1) To set up links between representations of practitions and the ATN parser-generator, so that CASSIE generates an utterance just when she intends to do so now—i.e. just when an intention to do so now is present in CASSIE. Analogously, the parser will be activated only when CASSIE intends to understand someone now. (A moving "now" pointer in SNePS is an outcome of the work by [2,3].) The resulting system can be expected to be "input/output cantankerous" and quite unlike existing systems—more like a human agent, in fact—since it will "decide" when to read input and when to produce output. (Trace facilities can be used to follow "internal" processes.) In addition, of course, this mechanism will be used to implement the understanding and generation of descriptions, as described in Section 2, above.

(2) To add a plan-based inferential mechanism, such that CASSIE will perform inferences in a truly natural deductive system (based on [17,43]; by "truly natural", we mean neither resolution-based, axiomatic, nor Gentzen-style sequents), adding not only the resulting conclusion to the network (as SNePS does now) but also any intermediate steps required by the "natural" reasoning, and storing the "plan" or strategy used in solving the problem so that it can be retrieved from memory and applied to future "similar" patterns. The plan-based inferential mechanisms, like the ATN parser-generator, will be under intentional control: CASSIE will reason when she intends to and when, in restricted scenarios, reasoning is automatically triggered.

5. SNePS and PSNePS.

The implementation of all parts of the project is being done using SNePS, implemented in Franz LISP and running on VAX 11/750s in the Department of Computer Science, SUNY Buffalo, and at SUNY Fredonia.

In addition, Dipert has been developing a Prolog program, PSNePS, that simulates part of SNePS and also provides an interface to translate any SNePS network built by this system into the SNePS User Language (SNePSUL), which is the standard interface to SNePS.

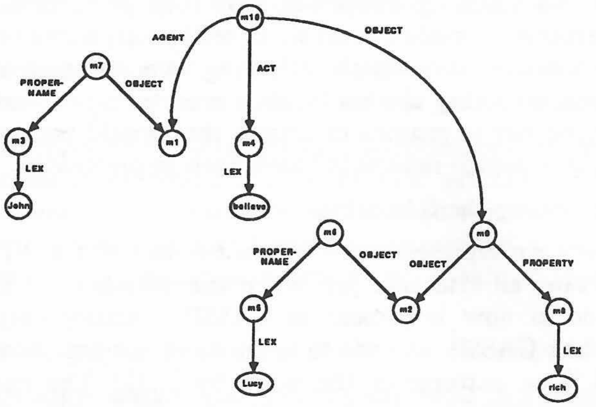


Figure 1 (1). A SNePS representation of the de re belief report 'John believes of Lucy that she is rich'.

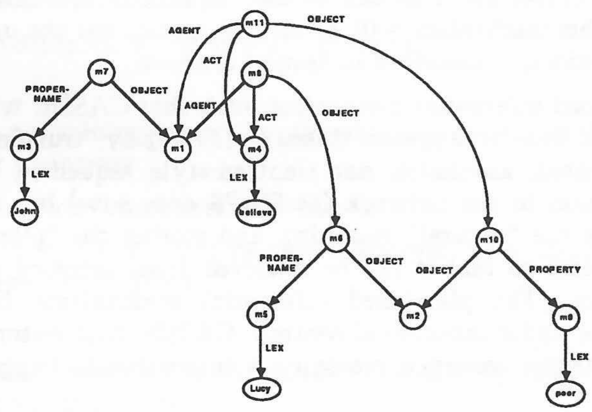


Figure 1 (2). A SNePS network for the de dicto belief report 'John believes that Lucy is poor'.

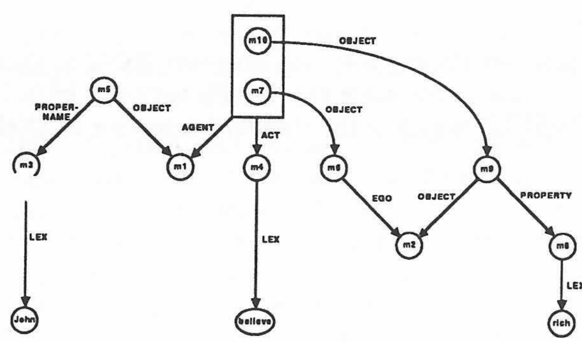


Figure 1 (3). SNePS network for 'John believes that he* is rich'.

(The box is a shorthand graphical representation to eliminate redundant AGENT and ACT arcs as a notational convenience.)

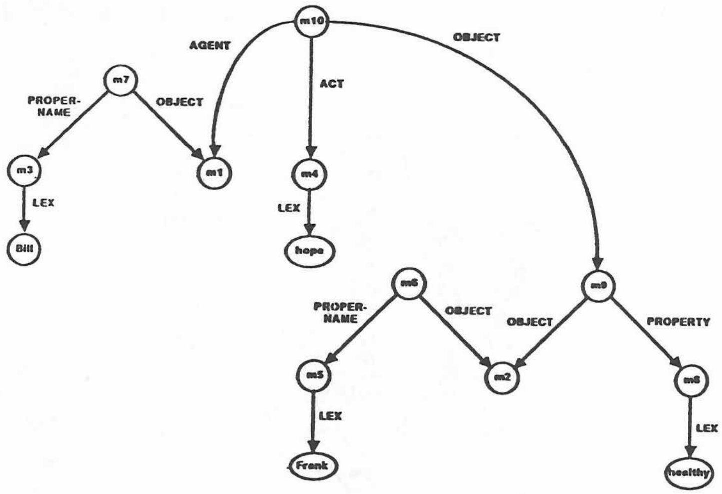


Figure 3a. A SNePS representation of the de re hope report 'Bill hopes of Frank that he is healthy'.

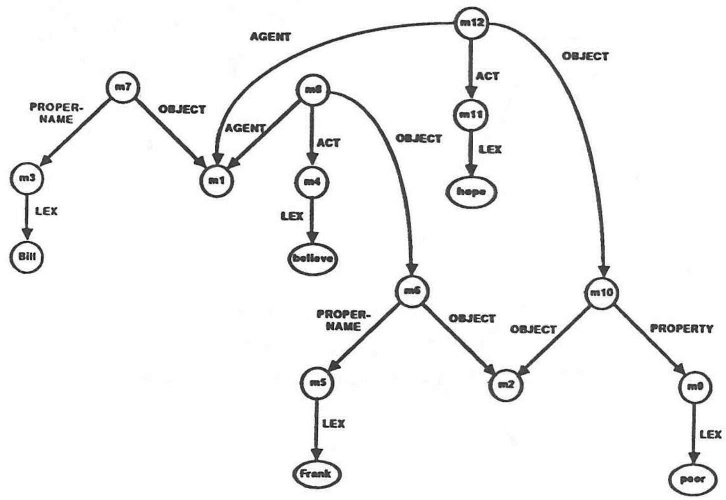


Figure 3b. A SNePS network for the de dicto hope report 'Bill hopes that Frank is healthy'.

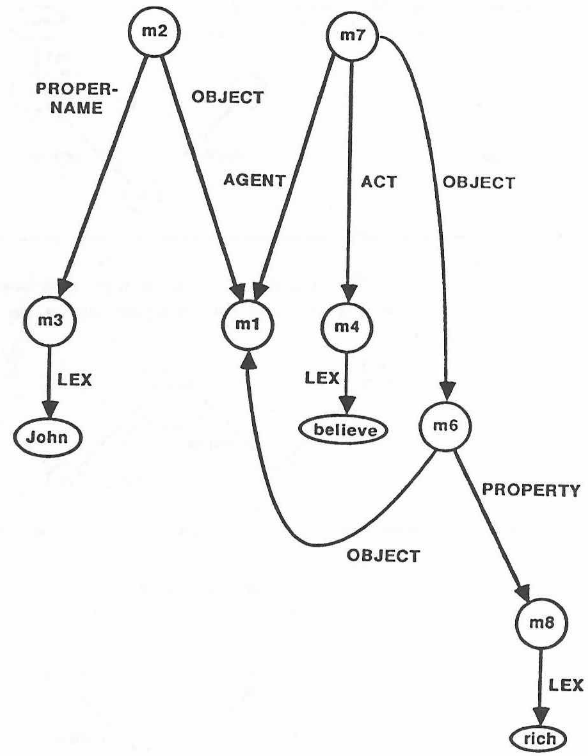


Figure 2 (3B). A SNePS network for 'John believes that he* is rich'.

Agent's own representation of:
"I ought$_{abs}$ I* to
drive a car."

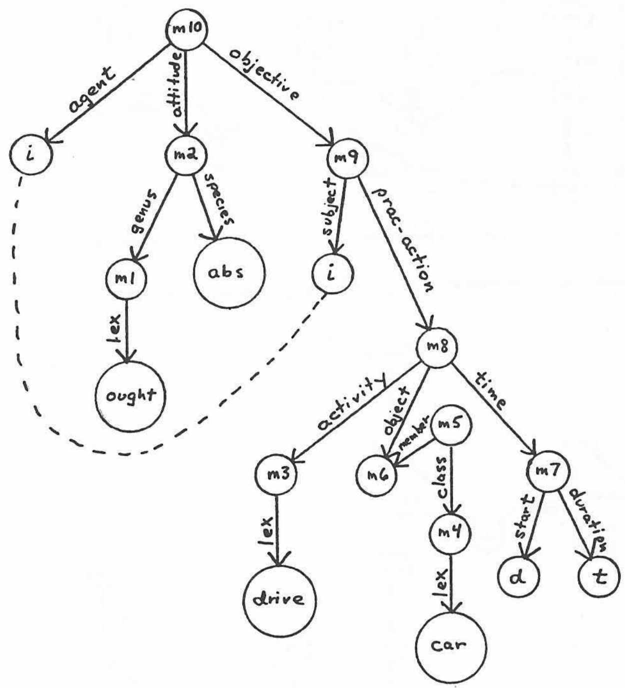


Figure 5 (1)

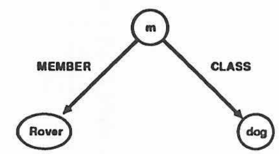


Figure 4a. A SNePS network for 'Rover is a dog'.

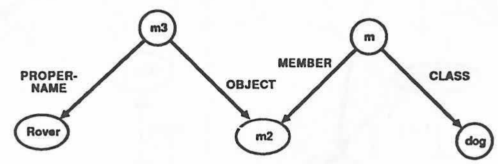


Figure 4b. Another SNePS network for 'Rover is a dog'.
(Here, node m2 is an unstructured individual.)

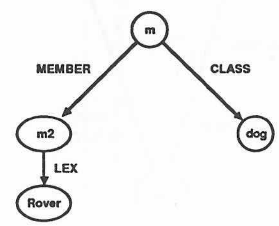


Figure 4c. A third SNePS network for 'Rover is a dog'.
(Here, node m2 is a structured individual.)

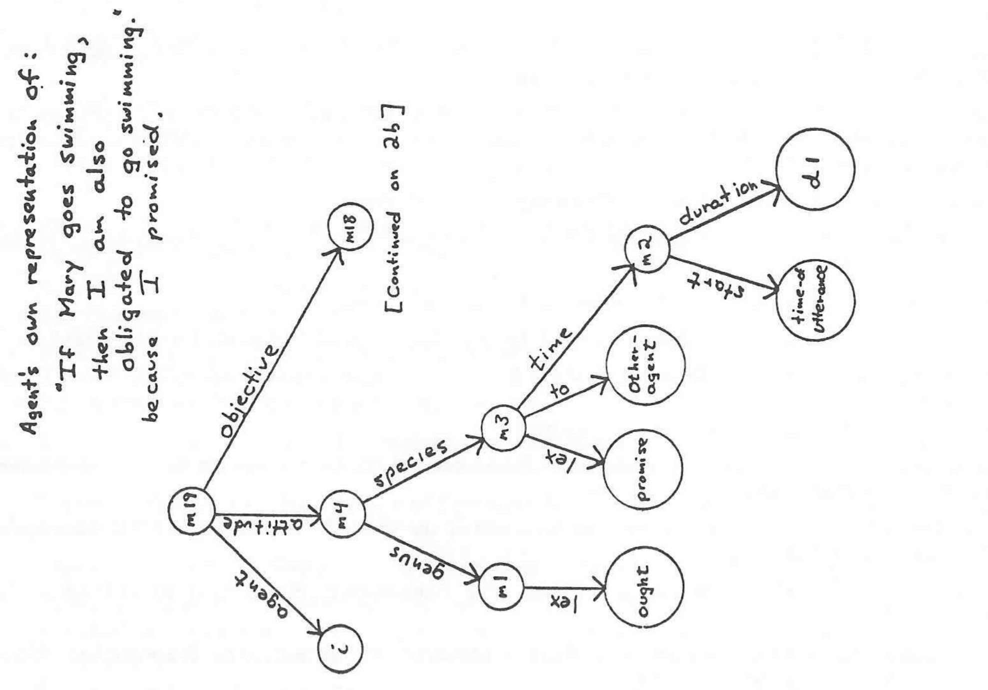


Figure 5(2a)

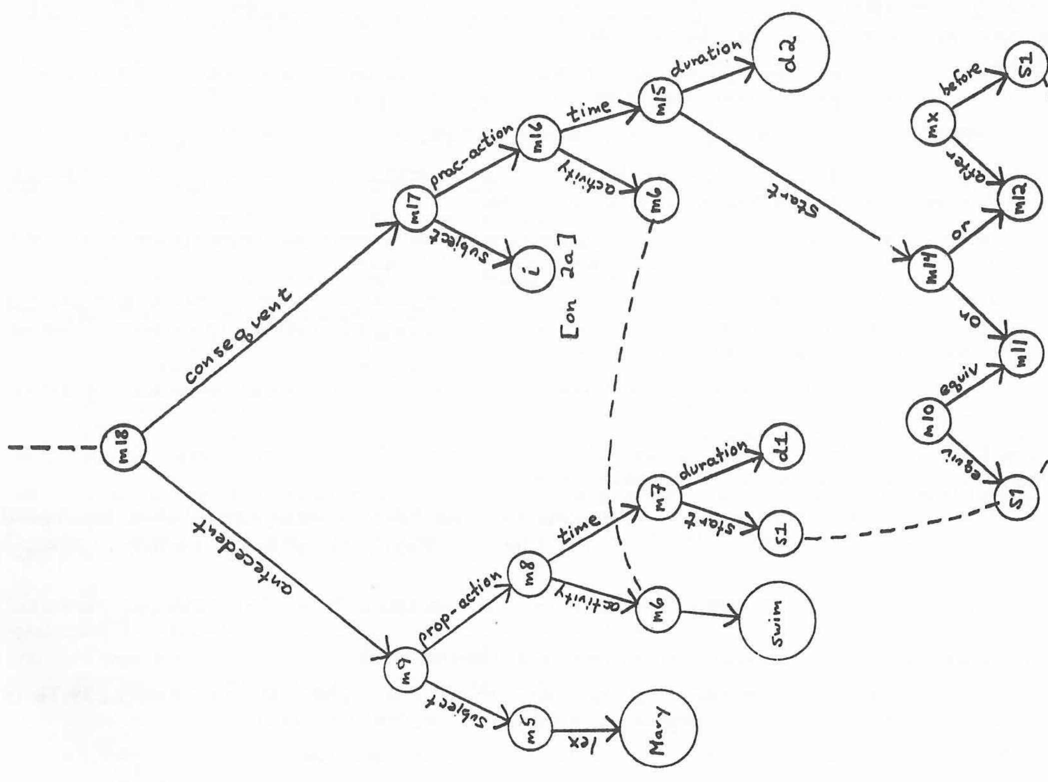


Figure 5(2b)

References

1. Allen, James F., & Perrault, C. Raymond (1980), "Analyzing Intentions in Utterance," *Artificial Intelligence* 15: 143-78.

2. Almeida, Michael J. (1987), "Reasoning about the Temporal Structure of Narratives," *Technical Report*, No. 87-10 (Buffalo: SUNY Buffalo Dept. of Computer Science).

3. Almeida, Michael J., & Shapiro, Stuart C. (1983), "Reasoning about the Temporal Structure of Narrative Texts," *Proc. 5th Annual Meeting of the Cognitive Science Society (University of Rochester)* (Hillsdale, NJ: Lawrence Erlbaum), unpaginated.

4. Brand, Myles (1984), *Intending and Acting* (Cambridge, MA: MIT Press).

5. Bratman, Michael (1986), "Intention and Commitment," paper presented at American Philosophical Association Pacific Division, Los Angeles.

6. Bratman, Michael (forthcoming), *Intention, Plans, and Practical Reason.*

7. Bratman, Michael, & Lansky, Amy (1986), "Rational Agency (RatAg)," *CSLI Monthly* 1.4 (June 1986).

8. Bruder, Gail A.; Duchan, Judith F.; Rapaport, William J.; Segal, Erwin M.; Shapiro, Stuart C.; & Zubin, David A. (1986), "Deictic Centers in Narrative: An Interdisciplinary Cognitive-Science Project," Technical Report No. 86-20 (Buffalo: SUNY Buffalo Dept. of Computer Science).

9. Castañeda, Hector-Neri (1970), "On the Philosophical Foundations of the Theory of Communication: Reference," *Midwest Studies in Philosophy* 2 (1977) 165-86.

10. 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.

11. Castañeda, Hector-Neri (1975), *Thinking and Doing: The Philosophical Foundations of Institutions* (Dordrecht: D. Reidel).

12. Christiansen, Alan D. (1985), "The History of Planning Methodology: An Annotated Bibliography," *SIGART Newsletter*, No. 84 (October 1985), pp. 44 ff.

13. Davidson, Donald (1980), "Intending," in *Essays on Actions and Events* (Oxford: Clarendon Press).

14. Davis, Randall (1980), "Meta-Rules: Reasoning about Control," *Artificial Intelligence* 15: 179-222.

15. Dipert, Randall R. (1986a), "Art, Artifacts, and Regarded Intentions," *American Philosophical Quarterly* 23: 401-408.

16. Dipert, Randall R. (1986b), "Intending and Acting" (paper presented to SNePS Research Group, SUNY Buffalo, and to Department of Philosophy, Indiana University).

17. Dipert, Randall R. (1986c), "The Epistemology of Proofs and Automated Theorem-Proving: What Is It to 'Understand' a Proof?" (paper presented to the SUNY Buffalo Logic Colloquium).

18. Donnellan, Keith S. (1966), "Reference and Definite Descriptions," *Philosophical Review* 60: 281-304.

19. Donnellan, Keith S. (1978), "Speaker Reference, Descriptions and Anaphora," in P. Cole (ed.), *Syntax and Semantics*, Vol. 9: Pragmatics (Orlando, FL: Academic Press): 47-68.

20. Fawcett, Brenda (1986), "The Representation of Ambiguity in Opaque Constructs," Technical Report No. CSRI-178 (Toronto: University of Toronto Computer Systems Research Institute).

21. Fawcett, Brenda, & Hirst, Graeme (1986), "The Detection and Representation of Ambiguities of Intension and Description," *Proc. 24th Annual Meeting Assoc. for Computational Linguistics (Columbia University)* (Morristown, NJ: Assoc. for Computational Linguistics): 192-99.

22. Grice, H. P. (1975), "Logic and Conversation," in P. Cole & J. Morgan (eds.), *Syntax and Semantics*, Vol. 3 (New York: Academic Press).

23. Konolige, Kurt (1984), "Belief and Incompleteness," in J. R. Hobbs & R. C. Moore (eds.), *Formal Theories of the Commonsense World* (Norwood, NJ: Ablex, 1985): 359-403.

24. Maida, Anthony S. (1986), "Introspection and Reasoning about the Beliefs of other Agents," *Proc. 8th Annual Conference of the Cognitive Science Soc. (Univ. of Mass. at Amherst)* (Hillsdale, NJ: Lawrence Erlbaum): 187-95.

25. Maida, Anthony S., & Shapiro, Stuart C. (1982), "Intensional Concepts in Propositional Semantic Networks," *Cognitive Science* 6: 291-330. Reprinted in R. J. Brachman and H. J. Levesque, eds. *Readings in Knowledge Representation* (Los Altos, CA: Morgan Kaufmann, 1985): 170-189.

26. McCarty, L. T. (1986), "Permissions and Obligations: An Informal Introduction," Technical Report LRP-TR-19 (New Brunswick, NJ: Laboratory for Computer Science Research, Rutgers University).

27. Parsons, Terence (1980), *Nonexistent Objects* (New Haven: Yale University Press).

28. Pollack, Martha E. (1986), "A Model of Plan Inference that Distinguishes between the Beliefs of Actors and Observers," *Proc. 24th Annual Meeting of the Assoc. for Computational Linguistics (Columbia University)* (Morristown, NJ: Assoc. for Computational Linguistics): 207-14.

29. Rapaport, William J. (1978), "Meinongian Theories and a Russellian Paradox," *Noûs* 12: 153-80; errata, *Noûs* 13 (1979) 125.

30. Rapaport, William J. (1979), "An Algebraic Interpretation of Deontic Logic" (abstract), *Journal of Symbolic Logic* 44: 472.

31. Rapaport, William J. (1984a), Critical Notice of Routley 1979, *Philosophy and Phenomenological Research* 44: 539-52.

32. Rapaport, William J. (1984b), "Belief Representation and Quasi-Indicators," Technical Report No. 215 (Buffalo: SUNY Buffalo Department of Computer Science).

33. Rapaport, William J. (1985a), "To Be and Not to Be: Critical Study of Parsons's *Nonexistent Objects*" *Noûs* 19: 255-71; erratum, *Noûs* 20 (1986) 587.

34. Rapaport, William J. (1985b), "Meinongian Semantics for Propositional Semantic Networks," *Proc. 23rd Annual Meeting Assoc. for Computational Linguistics (Univ. of Chicago)* (Morristown, NJ: Assoc. for Computational Linguistics): 43-48.

35. Rapaport, William J. (1985/1986), "Non-Existent Objects and Epistemological Ontology," *Grazer Philosophische Studien* 25/26: 61-95.

36. Rapaport, William J. (1986a), Review of Lambert's *Meinong and the Principle of Independence*, in *Journal of Symbolic Logic* 51: 248-52.

37. Rapaport, William J. (1986b), "Logical Foundations for Belief Representation," *Cognitive Science* 10: 371-422.

38. Rapaport, William J. (1986d), "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).

39. Rapaport, William J. (1987), "Belief Systems," in S. C. Shapiro (ed.), *Encyclopedia of Artificial Intelligence* (New York: John Wiley): 63-73.

40. Rapaport, William J., & Shapiro, Stuart C. (1984), "Quasi-Indexical Reference in Propositional Semantic Networks," *Proc. 10th Int'l. Conf. Computational Linguistics (COLING-84; Stanford Univ.)* (Morristown, NJ: Assoc. for Computational Linguistics): 65-70.

41. Rapaport, William J.; Shapiro, Stuart C.; & Wiebe, Janyce M. (1986), "Quasi-Indicators, Knowledge Reports, and Discourse," Technical Report No. 86-15 (Buffalo, NY: SUNY Buffalo Dept. of Computer Science).

42. Routley, Richard (1979), *Exploring Meinong's Jungle and Beyond* (Canberra: Australian National University, Research School of Social Sciences, Dept. of Philosophy).

43. Schagrin, Morton L.; Rapaport, William J.; & Dipert, Randall D. (1985) *Logic: A Computer Approach* (New York: McGraw-Hill).

44. Shapiro, Stuart C. (1979), "The SNePS Semantic Network Processing System," in N. V. Findler (ed.) *Associative Networks* (New York: Academic Press): 179-203.

45. Shapiro, Stuart C. (1982), "Generalized Augmented Transition Network Grammars For Generation From Semantic Networks," *American Journal of Computational Linguistics* 8: 12-25.

46. Shapiro, Stuart C., & Rapaport, William J. (1986), "SNePS Considered as a Fully Intensional Propositional Semantic Network," *Proc. 5th Nat'l. Conf. on Artificial Intelligence (AAAI-86; Philadelphia)* (Los Altos, CA: Morgan Kaufmann), Vol. 1, pp. 278-83.

47. Shapiro, Stuart C., & Rapaport, William J. (1987), "SNePS Considered as a Fully Intensional Propositional Semantic Network," in G. McCalla and N. Cercone (eds.), *The Knowledge Frontier* (New York: Springer-Verlag): 262-315.

48. 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.

49. Wilks, Yorick (1986), "Default Reasoning and Self-Knowledge," *Proceedings of the IEEE* 74: 1399-1404.

50. Wilks, Yorick, & Bien, Janusz (1983), "Beliefs, Points of View, and Multiple Environments," *Cognitive Science* 7: 95-119.

51. Woods, William A. (1975), "What's in a Link: Foundations for Semantic Networks," in D. G. Bobrow and A. M. Collins (eds.), *Representation and Understanding* (New York: Academic Press): 35-82.