1988-5

Quasi-Indexicals and Knowledge Reports

William J. Rapaport
Department of Computer Science, Department of Philosophy, and Center for Cognitive Science
State University of New York at Buffalo, Buffalo, NY 14260, U.S.A.
rapaport@cs.buffalo.edu

Stuart C. Shapiro
Department of Computer Science and Center for Cognitive Science
State University of New York at Buffalo, Buffalo, NY 14260, U.S.A.
shapiro@cs.buffalo.edu

Janyce M. Wiebe
Department of Computer Science and Computing Research Laboratory
New Mexico State University, Las Cruces, NM 88003, U.S.A.
wiebe@cs.nmsu.edu

12.1 Introduction.

How are knowledge and belief related? The standard philosophical analysis, dating back at least to Plato (*Theaetetus* 201), is that knowledge is justified true belief (but cf. Gettier 1963). In this paper, we describe some issues that are in the field of knowledge representation taken literally—issues in the representation of knowledge reports, where knowledge is treated as true belief. (Consideration of cognitive agents' *justifications* for their beliefs has not recently been of central concern to formal computational

F. Orilia and W.J. Rapaport (eds.), Thought, Language, and Ontology, 235–294.

analyses of knowledge (cf. Rapaport 1992 for a survey); however, once the appropriate logical foundations for knowledge- and belief-representation are determined, the issue of justification ought once again to become a major area of research.)

In particular, we present a computational analysis of *de re*, *de dicto*, and *de se* belief and knowledge reports. Our analysis solves a problem first observed by Hector-Neri Castañeda, namely, that the simple rule '(*A* knows that *P*) implies *P*' (the rule that "knowledge is true", so to speak) apparently does not hold if *P* contains a "quasi-indexical" (sometimes called a "quasi-indicator"). A *quasi-indexical* is an expression within an intentional context (e.g., a propositional-attitude context) that represents a use of an indexical by another person; indexicals, by contrast, make strictly demonstrative reference. In linguistics, something very like quasi-indexicals are studied under the term "logophoric pronouns". In addition, quasi-indexicals are often used as "shifted pronouns" in subjective contexts in narrative (in what is known as "represented thought" or "free indirect discourse"): For example, in "She [Hannah] winced as she heard them crash to the platform. The lovely little mirror that *she had brought* for Ellen, and the gifts for the baby!" (Franchere 1964: 3, emphasis added; cited in Wiebe & Rapaport 1988: 132), the second expression represents Hannah's thought upon hearing the crash. Note that were Hannah to have uttered this thought, she would have used the first person and simple past tense: "The lovely little mirror that *I brought*...". Note, too, that the second expression is neither a complete sentence nor capable of being embedded in a "Hannah thought that ..." context—it is what Ann Banfield (1982) calls an "unspeakable sentence". Further examples of quasi-indexicals will be given below. (Cf. Castañeda 1966, 1967; Hamburger 1973; Kuroda 1976; Banfield 1982; Rapaport & Shapiro 1984; Rapaport 1986; Wiebe & Rapaport 1986, 1988; Sells 1987; Roberts & Rapaport 1988; Wiebe 1990ab, 1991, 1994; Reboul 1992; Fludernik 1993.)

We present a single rule, in the context of a knowledge-representation and reasoning system, that holds that "knowledge is true" for all propositions *P*, including those whose expression in a "public communication language" such as English (Shapiro 1993) contains quasi-indexicals. (By 'rule', we mean an implication (i.e., an if–then proposition) that can be used as a recipe for drawing a conclusion based on evidence. It is *like* a rule in AI "rule-based" systems, though it is *not* a production rule. Nor is it a rule of inference, which is part of an inference engine and is not explicitly represented in a knowledge-representation system.) In so doing, we explore the difference between reasoning in a public communication language and in a knowledge-representation language, we demonstrate the importance of representing proper names explicitly, we provide support for the necessity of considering sentences in the context of extended discourse (for example, written narrative) in order to fully capture certain features of their semantics, and we reply to several objections to our earlier theory (Rapaport & Shapiro 1984, Rapaport 1986) raised by Yorick Wilks, Afzal Ballim, and Eric Dietrich (1989; cf. Ballim & Wilks 1991; for other commentary on Rapaport 1986, see Israel 1989 and Von Eckardt, Barbara 1993, §5.1). An appendix provides a transcript of sample runs of both old and revised computational implementations of our theory.

12.2 The Varieties of Belief.

At the very least, knowledge *implies* true belief and, thus, is a *kind* of belief. Now, among the kinds of belief reports, there are *de re*, *de dicto*, and *de se* belief reports.

12.2.1 *De Re* and *De Dicto* Belief Reports.

To understand the difference between *de re* and *de dicto* belief reports, consider the following claim:

(12.1) Columbus believed that Castro's island was India.

This claim can be understood in (at least) two ways. On one reading, it is true: Suppose you have asked me what body of land Columbus believed to be India. If I have temporarily forgotten the name 'Cuba', I might truthfully reply with (12.1) if we mutually know what body of land Castro's island is (namely, Cuba). This report of Columbus's belief is said to be *de re* (of the thing), since I use the noun phrase 'Castro's island' to refer to the body of land (the thing).

On another reading, (12.1) is false: Suppose you have asked me what Columbus believed when he sighted land. And suppose that I try to convey to you Columbus's subjective belief that the body of land he was then looking at was India. If I replied with (12.1), I would be misleading you, since Columbus would not have referred to that body of land as 'Castro's island' (nor as 'Cuba'; nor, for that matter, would he have described it as an island, especially if he had believed it to be India). Rather, I should say something like

(12.2) Columbus believed that the body of land that he saw before him was India.

Both this report of Columbus's belief (which is true, or at least more accurate) and my earlier, misleading (indeed, false) report of Columbus's belief are said to be *de dicto* (of what is said), since the noun phrases used to characterize the body of land employ words that Columbus might have used (in translation, of course!). They purport to show us the content of his thoughts.

Given a belief report, how does the hearer decide whether it should be interpreted *de re* or *de dicto*? This is a very difficult question. (For work addressing this problem in the special case of determining whether a narrative passage is a subjective context, see Wiebe & Rapaport 1988; Wiebe 1990ab, 1991, 1994.) In order to sidestep this issue for present purposes, we will resort to fairly standard canonical English expressions that allow belief reports to wear their *de re*-ness or *de dicto*-ness on their sleeves, so to speak. For example, (12.1) will be taken as the *de dicto* reading, and we will use

(12.3) Columbus believed of Castro's island that it was India.

for the *de re* reading. Note that, in the *de dicto* canonical expression, 'Castro's island' is within (the scope of) the 'that' clause, whereas, in the *de re* canonical expression, 'Castro's island' is outside the 'that' clause.

Thus, in general, we shall canonically express a *de re* belief report (made by a speaker *S* to a hearer *H*) as

(12.4) *A* believes of *N* that F.

This represents the claim (by *S*) that agent *A* believes that someone whom *S* (and possibly *H*) believes to be named (or described by) '*N*' has property F. (The single quotes around names are needed: *N* represents the person; '*N*' is the person's name.) As Castañeda (1970) puts it, such a report (at least in isolation) is *referentially transparent* but *propositionally opaque* (also dubbed 'relational' by Quine (1956/1976)). This means that '*N*' can be replaced by any co-referring expression, preserving truth value. That is, '*N*' is a "speaker's reference" and can be replaced by any expression that *S* believes is co-referential with it. Thus, the report is (referentially) transparent in that it lets us see through it directly to the referent. However, this is at the expense of omitting, or hiding, any information about *A*'s characterization of *N*. It is (propositionally) opaque in that it blocks us from seeing the content of the believer's belief. As we saw, from the *de re* report (12.3), we cannot infer that Columbus characterized the land he saw before him as being Castro's island.

We shall canonically express a *de dicto* belief report (made by speaker *S* to hearer *H*) as

(12.5) *A* believes that *N* is F.

This represents the claim (by *S*) that *A* believes that someone whom *she or he* (that is, *A*) believes to be named (or described by) '*N*' has property F. Such a report (again, at least in isolation) is *referentially opaque* but *propositionally transparent* (also dubbed 'notional' by Quine). That is, '*N*' is a "believer's reference", and can*not* be replaced by any expression that *S* believes is co-referential. For example, from the *de dicto* report,

(12.6) Columbus believed that Queen Isabella was interested in the New World,

we *can* infer that Columbus characterized her as "Queen Isabella", and we can*not* replace 'Queen Isabella' by, say, 'the woman described on page 1048 of the *Columbia Encyclopedia*' (even if those two noun phrases *are* co-referential). (The hedge "in isolation", here and above, serves to warn against complexities that arise in discourse, where many of these concepts break down; see Wiebe & Rapaport 1986.)

12.2.2 *De Se* Belief Reports.

We can now turn to *de se* belief reports. Consider, once again, the *de dicto* report (12.2), above. Since this is propositionally transparent, what *is* the proposition that we are supposing that Columbus believed? It is *not* (expressed by) the phrase following the 'that', namely, 'the body of land that *he saw* before *him was* India'. Rather, it would have been (in English translation, of course), 'the body of land that *I see* before *me is* India'. Note how the reporter of Columbus's belief has to *shift* from present tense to past ('see' → 'saw', 'is' → 'was') and from first person to third ('I' → 'he', 'me' → 'him'). This "deictic shift" (Hamburger 1973, Kuroda 1976, Banfield 1982, Bruder et al. 1986, Rapaport et al. 1989ab, Reboul 1992, Galbraith 1995, Segal 1995), required by English (and analogous transformations required in other languages; cf. Sells 1987), introduces *quasi*-indexicals into the *speaker's* report in order to represent the *believer's* indexicals. The quasi-indexical pronouns 'he' and 'him' (and the "quasi-indexical" verbs 'saw' and 'was'; cf. Castañeda 1989: 135–136) turn this *de dicto* report into a *de se* report—a report about a believer's beliefs about her- or himself.

Following Castañeda's notation, we mark quasi-indexicals with a '*'. Thus, we shall canonically express a *de se* belief report (made by speaker *S* to hearer *H*) as

(12.7) *A* believes that she* (or he*) is F.

This represents a *de dicto* report (by *S* to *H*) involving the quasi-indexical 'she*' (or 'he*'). Thus, (12.7) is the *reporter's* (*S*'s) way of expressing the *first-person belief* that *A* would express (using the indexical 'I') as: 'I am F'.

Our earlier work (Rapaport & Shapiro 1984, Rapaport 1986) has not been alone in noting the importance of quasi-indexicals in AI. Andrew R. Haas (1993; see §12.7.3, below) and Yves Lespérance and Hector J. Levesque (1995; cf. Lespérance 1989, Lespérance & Levesque 1990) also have theories recognizing their importance. In their theory of indexical knowledge and robot action, Lespérance and Levesque introduce two terms, "**self**, which denotes the current agent, and **now**, which denotes the current time" (1995: 80):

> When **self** occurs outside the scope of **Know** ..., it behaves like the English indexical "I", and when **now** occurs outside the scope of **Know** ..., it behaves like the English indexical "now". In the scope of **Know** on the other hand, **self** and **now** behave like quasi-indexicals—there are no temporal quasi-indexicals in English, but one can imagine how a temporal analog of "he himself" would work. (Lespérance & Levesque 1995: 82–83.)

However, as Castañeda has noted (1989; as we saw above), there *are* temporal quasi-indexicals in English. Furthermore, Lespérance and Levesque's "primitive indexicals" (as they call '**self**' and '**now**') are very much like their English counterparts: 'he', 'him', or 'himself', for instance, can be a pure (deictic) indexical when outside the scope of an intentional propositional attitude such as 'know' or 'believe' as well as a quasi-indexical when within its scope. For example, in:

John wrote himself a letter.

'himself' is a pure indexical, whereas in:

John believes that he (i.e., he himself) is rich.

'he', or 'he himself', is a quasi-indexical. Languages with logophoric pronouns, on the other hand, have distinct morphemes for use in quasi-indexical contexts.

Finally, note that Lespérance and Levesque are aware that certain formulas, such as universal instantiation ("Specialization", Proposition 3.1, p. 86) and introspection (Proposition 3.5: $\models \mathbf{Know}(\varphi) \supset \mathbf{Know}(\mathbf{Know}(\varphi))$, which has an implicit agent *a* as the knower; p. 87), require restrictions on the uses of '**self**' and '**now**'. But their version of the principle that "knowledge is true" (Proposition 3.4, Lespérance & Levesque 1995: 87),

$$\models \mathbf{Know}(\varphi) \supset \varphi$$

is silent about whether φ can contain '**self**' or '**now**'. If φ is of the form F(**self**), then Proposition 3.4 becomes: If *a* knows that **self** is F, then **self** is F. Since their '**self**' is only a *quasi*-indexical in the antecedent, but an *ordinary* indexical in the consequent, either they have our original problem (i.e., what is the pronominal antecedent of '**self**' as it occurs in the consequent?) or else they have a non-compositional semantics. Of course, they also note that their "logic is not intended to be a formalization of the behavior of English indexicals" (p. 82).

12.3 SNePS Representations of Belief Reports.

A computational agent capable of handling these reports in natural language has been implemented using a generalized augmented-transition-network (GATN) parser-generator connected to the SNePS knowledge-representation and reasoning system (Shapiro 1979b, 1982, 1990; Rapaport & Shapiro 1984; Rapaport 1986; Shapiro & Rapaport 1987, 1992, 1995; Shapiro & Martins 1990; Shapiro & Group 1994). Although the knowledge-representation and reasoning formalism we are using in this paper is SNePS, the conclusions we reach are applicable in general.

12.3.1 Syntax and Semantics.

In this section, we present a formal syntax and semantics for the fragment of SNePS used in this paper. As the syntax and semantics will make apparent, this fragment is a reified, first-order predicate calculus representation. There are many aspects of SNePS that are not discussed here, since they are not directly relevant to this paper. For instance, SNePS has many more rules of inference than the one given below. (For more details on SNePS, see the references cited above.)

12.3.1.1 Syntax.

In this section, and subsequently, meta-linguistic variables are set in *italic* font, object-language symbols are set in `typewriter` font, and variables that range over object-language symbols are set in a lower-case sans serif font.

Relations

Some set of Lisp symbols are distinguished as the *relations*. In this paper, this set is {`ego`, `lex`, `equiv`, `object`, `propername`, `property`, `agent`, `act`, `forall`, `ant`, `cq`}.

Case Frames

Some subset of the powerset of relations is distinguished as the set of *case frames*. In this paper, this set is { {`ego`}, {`lex`}, {`equiv`}, {`object`, `propername`}, {`object`, `property`}, {`agent`, `act`, `object`}, {`forall`, `ant`, `cq`} }.

Atomic Terms

1. If *w* is a lexeme, then w is a *lexical atomic term*. Examples of lexical atomic terms used in this paper are `John`, `poor`, and `rich`.
2. If *n* is an integer, then `B`n is a *base atomic term*. Examples of base atomic terms used in this paper are `B1`, `B2`, and `B3`.
3. If *n* is an integer, then `V`n is a *variable atomic term*, or, simply, a *variable*. Examples of variables used in this paper are `V1` and `V2`.

In graphical diagrams, an atomic term is shown as the term inside a circle or oval.

Well-Formed Terms

1. If t is an atomic term, then t is a well-formed term (wft).
2. If $\{r_1, \ldots, r_n\}$ is a case frame, and $\{t_1, \ldots, t_n\}$ are wfts, then the set of pairs $\{\langle r_1, t_1\rangle, \ldots, \langle r_n, t_n\rangle\}$ is a (molecular) wft. The wfts $t_1, \ldots, t_n$ are said to be *immediately dominated* by the molecular wft. If there is a sequence of terms, $t_1, \ldots, t_n$, where each t_i immediately dominates t_{i+1}, then t_1 is said to *dominate* t_n.

   A pair in a molecular wft may also consist of a relation and a *set* of wfts (in which case, each wft in the set is immediately dominated by the molecular wft). In this paper, however, this option will only be used when the relation of the pair is either `equiv` or `forall`. For ease of presentation, we will assume that `equiv` and `forall` are always paired with sets of wfts, even if they are singleton sets.

   For ease of reference, each molecular wft is assigned an *identifier*, either of the form `M`n or of the form `P`n, where n is an integer. To show the assignment of an identifier to a wft, the identifier is written before the wft with a colon (':') between them. For example,

`M10:{⟨object, B3⟩, ⟨property, M9:{⟨lex, rich⟩}⟩}`

is a wft. Once the assignment is known, the identifier may be used on its own, for example,

`M11:{⟨agent, B2⟩, ⟨act, M1⟩, ⟨object, M10⟩}.`

In graphical diagrams, a molecular wft is shown as its identifier inside a circle or oval, and, for each pair ⟨r, t⟩ in the wft, a directed arc is drawn from the diagram of the molecular wft to the diagram of t, labeled with the relation r.

Open and Closed Terms

If v is a variable, and m is a molecular wft containing the pair ⟨`forall`, {..., v, ...}⟩, then we say that v is *bound in* m. If a molecular wft m dominates a variable v, and v is neither bound in m nor in any other wft dominated by m, then v is said to be *free in* m. (In this paper, we will not be concerned with the possibility that a wft m could contain some bound and some free occurrences of the same variable v.) We also say that v is free in the term v itself. A wft with at least one variable free in it is *open;* otherwise, it is *closed.* Open wfts are assigned identifiers of the form Pn; closed wfts are assigned identifiers of the form Mn.

Asserted Terms

Wfts are either *asserted* or *unasserted.* An asserted wft is written with an exclamation mark ("!") following either the term itself (e.g., see (RI), below), or its identifier when that is written. For example, writing `M11!:{⟨agent, B2⟩, ⟨act, M1:{⟨lex, know⟩}⟩, ⟨object, M10!⟩}` indicates that both `M11!` and `M10!` are asserted. Only closed wfts may be asserted.

In graphical diagrams, an asserted wft is shown with an exclamation mark following its identifier.

Substitution Instances

If $\{v_1, \ldots, v_n\}$ are variables that occur free in a wft m, and $\{t_1, \ldots, t_n\}$ are wfts, then $m\{t_1/v_1, \ldots, t_n/v_n\}$ is the wft that results from copying m but replacing every occurrence of each v_i by the corresponding t_i. We say that $m\{t_1/v_1, \ldots, t_n/v_n\}$ is an *instance* of m, and we call $\{t_1/v_1, \ldots, t_n/v_n\}$ a *substitution.*

Networks

A collection of wfts will be called a *network.*

Rule of Inference

$$\text{(RI)} \quad \frac{\mathtt{m}_1\sigma!,\ \{\langle\mathtt{forall},\ \mathtt{v}\rangle,\ \langle\mathtt{ant},\ \mathtt{m}_1\rangle,\ \langle\mathtt{cq},\ \mathtt{m}_2\rangle\}!}{\mathtt{m}_2\sigma!}$$

That is, if a network N contains an asserted wft of the form $\{\langle\mathtt{forall},\ \mathtt{v}\rangle,\ \langle\mathtt{ant},\ \mathtt{m}_1\rangle,\ \langle\mathtt{cq},\ \mathtt{m}_2\rangle\}$, and there is a substitution σ such that $\mathtt{m}_1\sigma$ is an asserted wft in N, then $\mathtt{m}_2\sigma$ may be added as an asserted wft in N. If $\mathtt{m}_2\sigma$ is already a wft in N, but unasserted, it may be changed to be asserted. (As noted earlier, full SNePS has more rules of inference; this is the only one we need in the subset of the logic that we are presenting. But this is not a significant limitation: (RI) is a form of resolution; i.e., it is a combination of universal instantiation (i.e., universal-quantifier elimination) and Modus Ponens, which are are typically the only rules of inference in first-order logic.)

12.3.1.2 Semantics.

SNePS semantics are given in terms of an agent A, a domain of discourse D, and a network N. Whenever A first conceives of an entity e of D, she does so by constructing a term t in N. We say that t *represents* e for A, and write $[\![\mathtt{t}]\!]_A = e$, leaving off the subscript when it is obvious from context. When e is a belief report, t represents what A understands when understanding e.

It is significant that an agent's network is constructed over time (as in Discourse Representation Theory; cf. Kamp & Reyle 1993). Additional wfts may be added to an agent's network when she reads or is told something, or when she uses (RI) to infer something.

The Agent

A SNePS network is considered to be the mind of a cognitive agent. For ease of exposition, when we consider a single agent, we will call her Cassie. If we need to talk about a second agent, we will call him Oscar.[1] Oscar's and Cassie's minds are independent networks: There is no necessary connection between a wft in one and wfts in the other, even those that look identical; the assignment of identifiers to wfts is done in the two networks independently.

The Domain of Discourse

The domain of discourse consists of all entities of the commonsense world that a cognitive agent could think about, read about, hear about, sense, or discuss, including real objects, fictional entities, and impossible entities; and including the agent itself, other agents, properties, and propositions—in short, "intensional"

[1] Cassie is the *C*ognitive *A*gent of the *S*NePS *S*ystem—an *I*ntelligent *E*ntity, first introduced in Shapiro & Rapaport 1987. Oscar is the *O*ther *S*NePS *C*ognitive *A*gent *R*epresentation, first introduced in Rapaport, Shapiro, & Wiebe 1986.

or "Meinongian" entities in "Aussersein", the domain of the objects of thought (see Meinong 1904; Rapaport 1978, 1981, 1985, 1985/1986, 1991ab; Maida & Shapiro 1982; Shapiro & Rapaport 1987, 1991; and cf. Hobbs's (1985) "Platonic universe"). We assume that different agents can share a single domain of discourse. So it is possible for Cassie's network to contain a wft t that represents the same entity that is represented for Oscar by Oscar's term t'. That is, $[\![t]\!]_{Cassie} = [\![t']\!]_{Oscar}$.

Note that The Morning Star and The Evening Star are different entities in the domain of discourse, even if they denote the same real-world object. When we want to say that two entities, e_1 and e_2, denote the same real-world object, we will write $e_1 \equiv e_2$.

Asserted Terms

The asserted terms of an agent A's network represent the beliefs of A. (Since forgetting and belief revision are independent of the arguments of this paper, they are ignored here, but see Martins & Shapiro 1988, Chalupsky 1993.)

Atomic Terms

1. A lexical atomic node w represents the lexeme *w* as a spoken, written, heard, or read token. For example, the node `poor` represents 'poor' and the node `John` represents 'John'.
2. A base atomic node `Bn` represents an entity that, at the time the agent constructs it, is not recognized by the agent as being identical to any entity already represented by a wft in the network.

   For example, if Cassie hears for the first time about something called 'The Morning Star', she might build a base node and assert of it that it is called 'The Morning Star'. Hearing later of something called 'The Evening Star' but not knowing that The Morning Star *is* The Evening Star (or so the standard philosophical example goes), she would build *another* base node and assert of it that it is called 'The Evening Star'.

   As another example, suppose Oscar says to Cassie, "The author of *Waverley* is in the next room." If Cassie knows that Scott is the author of *Waverley*, then she would just use her base node for Scott to represent the author of *Waverley*. If she did *not* know that Scott was the author of *Waverley*, then she would need a *new* base node to represent the author of *Waverley*.

   Or consider "A stockbroker who knows Bill likes him" (Kamp & Reyle 1993: 88). In SNePS, unlike Discourse Representation Theory, there is no need to create a new base node (or "discourse referent") for 'him', since the reader recognizes "him" as an entity, viz., Bill, already represented.

Well-Formed Terms

The semantics of molecular wfts are given individually for each case frame.

1. If w is a lexical atomic node, then $[\![\{\langle \texttt{lex}, \ \texttt{w}\rangle\}]\!]$ is the entity expressed by *w* when talking with another agent. For example, $[\![\{\langle \texttt{lex}, \ \texttt{poor}\rangle\}]\!]$ is the property referred to as 'poor' in English.

2. If n is a lexical atomic node, then $[\![\{\langle \texttt{object}, \ \texttt{o}\rangle, \ \langle \texttt{propername}, \ \texttt{n}\rangle\}]\!]$ is the proposition that the proper name of entity $[\![\texttt{o}]\!]$ is $[\![\texttt{n}]\!]$. For example, $[\![\{\langle \texttt{object}, \ \texttt{B2}\rangle, \ \langle \texttt{propername}, \ \texttt{John}\rangle\}]\!]$ is the proposition that $[\![\texttt{B2}]\!]$'s name is 'John'.

3. $[\![\{\langle \texttt{object}, \ \texttt{o}\rangle, \ \langle \texttt{property}, \ \texttt{p}\rangle\}]\!]$ is the proposition that entity $[\![\texttt{o}]\!]$ has the property $[\![\texttt{p}]\!]$. For example, $[\![\{\langle \texttt{object}, \ \texttt{B2}\rangle, \ \langle \texttt{property}, \ \{\langle \texttt{lex}, \ \texttt{rich}\rangle\}\rangle\}]\!]$ is the proposition that $[\![\texttt{B2}]\!]$ is rich.

4. $[\![\{\langle \texttt{agent}, \ \texttt{g}\rangle, \ \langle \texttt{act}, \ \texttt{a}\rangle, \ \langle \texttt{object}, \ \texttt{o}\rangle\}]\!]$ is the proposition that agent $[\![\texttt{g}]\!]$ performs the act $[\![\texttt{a}]\!]$ on object $[\![\texttt{o}]\!]$. As in Rapaport 1986, we will use this case frame in this paper even for propositional attitudes (sometimes called "mental acts"; cf. Rapaport 1978). For example, $[\![\{\langle \texttt{agent}, \ \texttt{B2}\rangle, \ \langle \texttt{act}, \ \{\langle \texttt{lex}, \ \texttt{know}\rangle\}\rangle, \ \langle \texttt{object}, \ \texttt{M7}\rangle\}]\!]$ is the proposition that $[\![\texttt{B2}]\!]$ knows $[\![\texttt{M7}]\!]$.

5. $[\![\texttt{m}:\{\langle \texttt{ego}, \ \texttt{n}\rangle\}]\!]$ is the proposition that $[\![\texttt{n}]\!]$ is the agent who believes $[\![\texttt{m}]\!]$. It is the thesis of this paper that the semantics of this case frame, introduced in Rapaport & Shapiro 1984 and Rapaport 1986, is actually incoherent.

6. $[\![\{\langle \texttt{equiv}, \ \{\texttt{t}_1, \ldots, \texttt{t}_n\}\rangle\}]\!]$ is the proposition that, for all $1 \leq i, j \leq n$, $[\![\texttt{t}_i]\!] \equiv [\![\texttt{t}_j]\!]$.

7. If $\texttt{m}:\{\langle \texttt{forall}, \ \texttt{v}\rangle, \ \langle \texttt{ant}, \ \texttt{a}\rangle, \ \langle \texttt{cq}, \ \texttt{c}\rangle\}$ is closed, then $[\![\texttt{m}]\!]$ is the proposition that, for every substitution σ, whenever $[\![\texttt{a}\sigma]\!]$ is believed by the agent, the agent may also believe $[\![\texttt{c}\sigma]\!]$.

12.3.2 Linear Representation of SNePS Networks.

Given this formal language, the formal *linear* SNePS representations of *de re*, *de dicto*, and *de se* belief reports (as given in Rapaport 1986) are given in (F1)–(F3), below. More precisely, these are representations *of what Cassie understands* when she understands the belief report, but we will occasionally talk loosely of "representations of belief reports". In (F1)–(F3), the "principal" `agent-act-object` node shown would be asserted (i.e., marked with an '!') in the case that Cassie actually believed it. The formal *graphical* SNePS representations of reports (F1)–(F3) are shown in Figures 12.1–12.3, respectively.

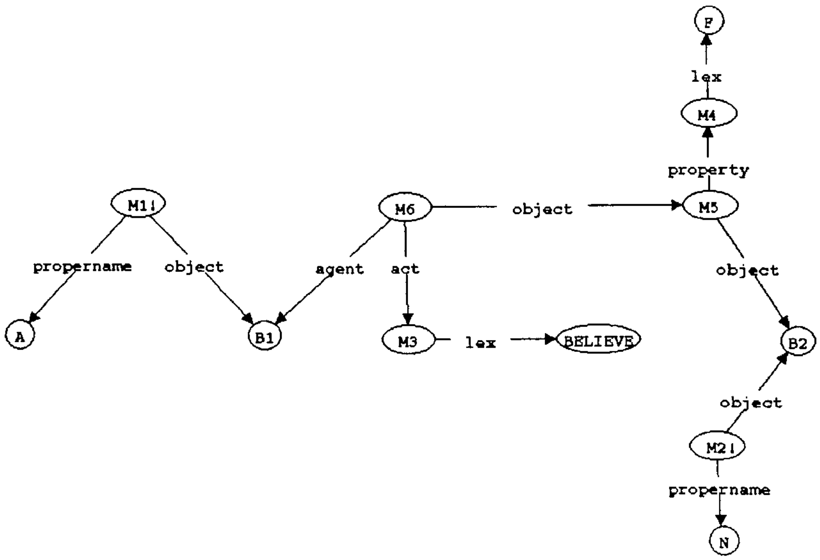

Figure 12.1: *A* believes (*de re*) of *N* that F. As an informal gloss to aid in reading this network:
M1!: (Cassie's belief that) B1's name is '*A*';
M2!: (Cassie's belief that) B2's name is '*N*';
M5: B2 is (i.e., has the property) F;
M6: B1 believes M5.

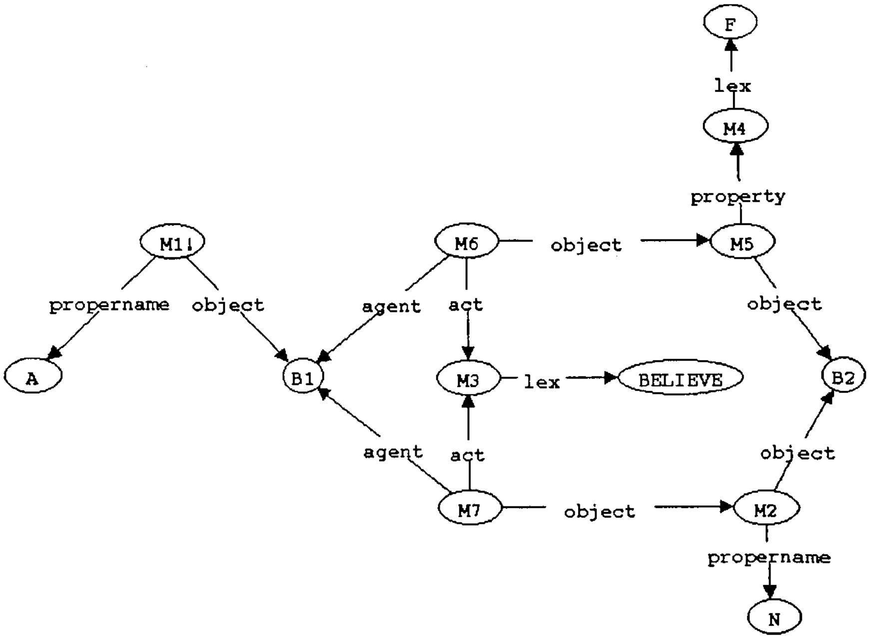

Figure 12.2: *A* believes (*de dicto*) that *N* is F:
M1!: (Cassie's belief that) B1's name is '*A*';
M2: B2's name is '*N*';
M5: B2 is F;
M6: B1 believes M5;
M7: B1 believes M2.

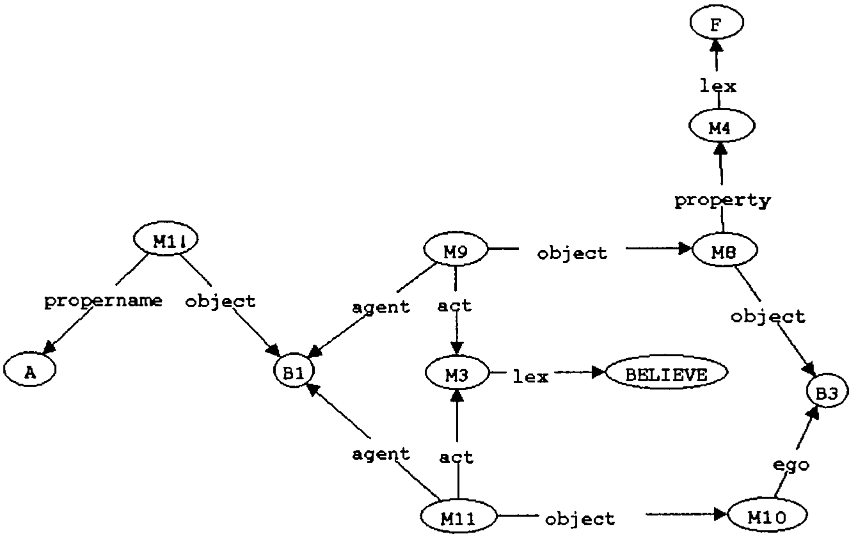

Figure 12.3: *A* believes (*de dicto*) that s/he* is F. (A *de se* belief report.)
M1!: (Cassie's belief that) B1's name is '*A*';
M8: B3 is F;
M9: B1 believes M8;
M10: "B3 is me" (see text);
M11: B1 believes M10.

(F1) The *de re* belief report: *A* believes of *N* that F:

```
M6:{⟨agent, B1⟩,
    ⟨act, M3:{⟨lex, believe⟩}⟩,
    ⟨object, M5⟩},
```

where:

```
M1!:{⟨object, B1⟩, ⟨propername, A⟩},
M2!:{⟨object, B2⟩, ⟨propername, N⟩},
M5:{⟨object, B2⟩, ⟨property, M4:{⟨lex, F⟩}⟩}
```

(F2) The *de dicto* belief report: *A* believes that *N* is F:

```
M6:{⟨agent, B1⟩,
    ⟨act, M3:{⟨lex, believe⟩}⟩,
    ⟨object, M5⟩},
```

where `M1!` and `M5` are as in (F1), `M2` is as in (F1), except that it need not be asserted, and:

```
M7:{⟨agent, B1⟩, ⟨act, M3⟩, ⟨object, M2⟩}
```

(F3) The *de se* belief report: *A* believes that she* (or he*) is F:

```
M9:{⟨agent, B1⟩, ⟨act, M3⟩, ⟨object, M8⟩},
```

where `M1!` and `M3` are as in (F1) and (F2), and:

```
M8:{⟨object, B3⟩, ⟨property, M4⟩},
```

where `M4` is as in (F1) and (F2), and:

```
M11:{⟨agent, B1⟩, ⟨act, M3⟩, ⟨object, M10:{⟨ego, B3⟩}⟩}
```

The semantic interpretations of (F1)–(F3) are as follows:

(FS1) $[\![$M6$]\!]$ is the proposition that *A* (i.e., $[\![$B1$]\!]$) believes $[\![$M5$]\!]$;
$[\![$M1$]\!]$ is the proposition that $[\![$B1$]\!]$'s name is '*A*';
$[\![$M2$]\!]$ is the proposition that $[\![$B2$]\!]$'s name is '*N*';
$[\![$M5$]\!]$ is the proposition that $[\![$B2$]\!]$ has property F.

(FS2) $[\![$M6$]\!]$ is the proposition that *A* believes $[\![$M5$]\!]$;
$[\![$M7$]\!]$ is the proposition that *A* believes $[\![$M2$]\!]$.

(FS3) $[\![$M9$]\!]$ is the proposition that *A* believes $[\![$M8$]\!]$;
$[\![$M8$]\!]$ is the proposition that $[\![$B3$]\!]$ has property F;
$[\![$M11$]\!]$ is the proposition that *A* believes that she* (or he*) is $[\![$B3$]\!]$.

For readers who find this precise formalism a bit hard to read, we can give *informal* analogues of the formal SNePS networks in a more familiar-looking (higher-order) predicate logic (with a "pretend-it's-English" semantics), as follows:

(I1) Propername(B1, '*A*') & Propername(B2, '*N*') & Believe(B1, F(B2))

(I2) Propername(B1, '*A*') & Believe(B1, Propername(B1, '*N*'))
& Believe(B1, F(B2))

(I3) Propername(B1, '*A*') & Believe(B1, Ego(B3)) & Believe(B1, F(B3))

The *de re* (I1)—corresponding to (F1) and Figure 12.1—says that ⟦B1⟧ is named '*A*', ⟦B2⟧ is named '*N*', and ⟦B1⟧ believes of ⟦B2⟧ that F. The *de dicto* (I2)—corresponding to (F2) and Figure 2—says that ⟦B1⟧ is named '*A*', ⟦B1⟧ believes of ⟦B2⟧ that ⟦B2⟧ is named '*N*', and ⟦B1⟧ believes of ⟦B2⟧ that F. Note that *de dicto* reports are analyzed in terms of *two de re* reports linked via the common term B2. (For an alternative to our representation of *de dicto* belief reports, see Wyatt 1989, 1990; cf. Wyatt 1993.) Finally, the *de se* (I3)—corresponding to (F3) and Figure 12.3—says that ⟦B1⟧ is named '*A*', ⟦B1⟧ believes of ⟦B3⟧ that ⟦B3⟧ is herself* (or himself*) (that is, 'Ego(B3)' is the proposition that ⟦B1⟧ would express as, roughly, '⟦B3⟧ is me'), and ⟦B1⟧ believes of ⟦B3⟧ (thus, of herself* (or himself*)) that F. Note that (I3) (and (F3) and Figure 12.3) represents a *de dicto* report.

In the remainder of the paper, we shall use the formal graphical representations as well as the informal linear notation. Readers may refer to whichever they find easier to read, though it should be kept in mind that only the *formal* (linear or graphical) notations are "official".

12.4 What Is Known Is True.

Now, just as there are *de re*, *de dicto*, and *de se* belief reports, so, it would seem, there ought to be *de re*, *de dicto*, and *de se* knowledge reports. In the rest of the paper, we shall consider to what extent this is so, how various knowledge reports are logically related to their corresponding or underlying belief reports, and the crucial role that extended discourse (such as written narrative) plays in the analysis.

Since knowledge is *true* belief, epistemic logics (cf. Hintikka 1962; Halpern 1986; Rapaport 1988a, 1992) have what Barwise and Perry (1983: 196) call "veridicality" as a thesis:

(V) $(A \text{ knows that } P) \rightarrow P$

where '*A*' names a cognitive agent and *P* is a proposition. We shall say that *P* is the proposition that is the *objective* of *A*'s mental act of knowing, using the Meinongian terminology of Rapaport 1985 and Shapiro & Rapaport 1987. We might express this rule in our system as shown in Figure 12.4; node M13 represents the rule whose antecedent is represented by node P1 (roughly, B1 knows V1) and whose consequent is represented by node V1 (corresponding to proposition *P*). In our informal linear notation, this would be:

Propername(B1, '*A*') & $\forall P$[Know(B1, P) $\rightarrow P$]

Prima facie, however, there are three problems with this. First, is such a rule even needed in the system? At first sight, it does not seem to be necessary, since if the system

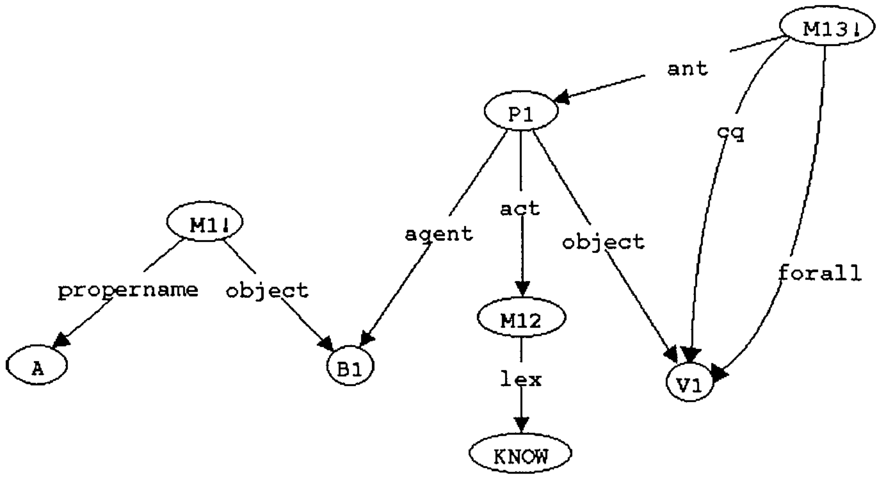

Figure 12.4: (A knows that P) $\rightarrow P$:
P1: B1 knows V1;
M13!: (Cassie's belief that) $\forall$V1[P1 $\rightarrow$ V1].

believes that A knows that P, then surely the system *already* believes that P and, hence, does not have to *infer* it. But suppose the system comes to believe that A knows that P because a highly reliable source told it so. It could, then, come to believe that P by inferring it, using (V). So let us assume that (V) should be in the system.

The next two problems that we must face are these: (1) Is (V) correct for both *de re* and *de dicto* knowledge reports? (2) Is (V) correct for *de se* knowledge reports involving quasi-indexicals? Let us agree to the following canonical expressions for *knowledge* reports (which are analogous to the canonical forms for belief reports, given in §12.2.1):

(12.8) A knows that N is F

will express a *de dicto* knowledge report, to be understood (given that knowledge implies belief) as implying that A believes (*de dicto*) that N is F; and

(12.9) A knows of N that F

will express a *de re* knowledge report, to be understood as implying that A believes (*de re*) of N that F.

12.5 *De Re* and *De Dicto* Knowledge.

Is (V) correct for both *de re* and *de dicto* knowledge reports? We can split (V) into two rules, corresponding to the two kinds of reports:

(V.dd) (*A* knows that *N* is F) → (*N* is F)

(V.dr) (*A* knows of *N* that F) → (*N* is F)

To express these rules more precisely, we must realize that the belief reports and other propositions represented in the system should be treated as beliefs of the system. That is, the system should be treated as a cognitive agent.

Thus, in the *de dicto* case, we can express the thesis that knowledge is true belief as follows (recall that we use the name 'Cassie' for the cognitive agent implemented by our system):

(KTB.dd)
(Cassie believes that *A* knows that *N* is F)
→ (Cassie believes that *A* believes that *N* is F) & (Cassie believes that *N* is F)

Note that now it is no longer the case that knowledge implies true belief *simpliciter*. Rather, Cassie's belief about a knowledge report (the antecedent of (KTB.dd)) implies two things: It implies her belief about a belief report (the first conjunct of the consequent of (KTB.dd)), and it implies her belief about the objective of that report (the second conjunct). The consequent of (KTB.dd) trivially implies (by &-elimination) that Cassie believes that *N* is F, agreeing with (V.dd).

In the *de re* case, the knowledge-is-true-belief thesis becomes:

(KTB.dr)
(Cassie believes that *A* knows of *N* that F)
→ (Cassie believes that *A* believes of *N* that F) & (Cassie believes that *N* is F)

whose consequent again trivially implies that Cassie believes that *N* is F, agreeing with (V.dr).

Before proceeding, it is important to be clear about a central point. In the *de dicto* case, (KTB.dd), '*N*' is *A*'s characterization of the individual that is F; that is, *N* must be "in" *A*'s "belief space". (For a definition of 'belief space', see §12.8.) In the *de re* case, (KTB.dr), '*N*' is *Cassie's* characterization of the individual that is F; that is, *N* must be directly in *Cassie's* belief space, but is not necessarily in *A*'s. Of course, all nodes are trivially in Cassie's belief space. For example, in the *de dicto* case, '*N*' is really Cassie's characterization of *A*'s characterization of the individual. Thus, in a sense, *all* of Cassie's beliefs "are *de dicto*"; i.e., the network that represents Cassie's beliefs is a *de dicto* representation: Her thoughts are propositionally transparent to us.

The SNePS representation of the *de re* rule is shown in Figure 12.5. In our informal predicate notation, it is:

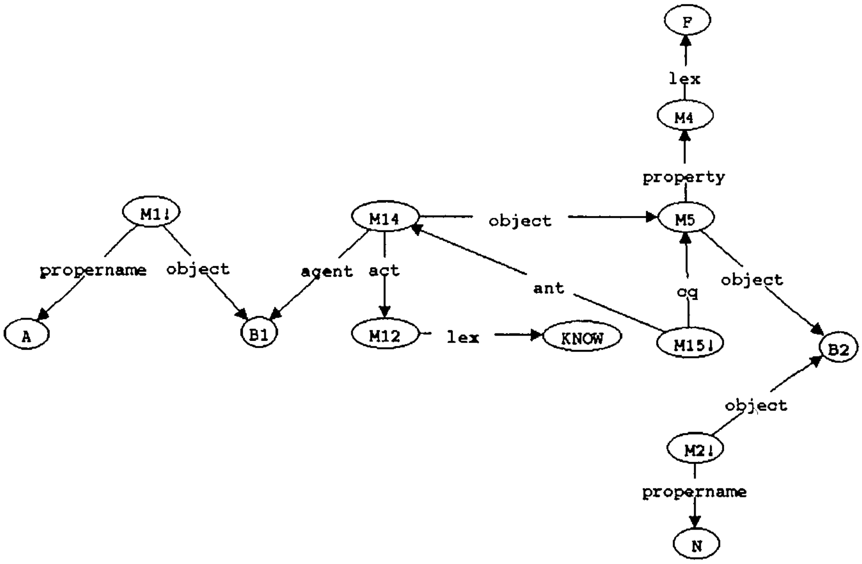

Figure 12.5: (A knows of N that F) $\rightarrow$ N is F.
M14: B1 knows M5;
M15!: (Cassie's belief that) M14 implies M5.

(F5) $$\text{Know}(\texttt{B1}, \text{F}(\texttt{B2})) \rightarrow \text{F}(\texttt{B2})$$

where Propername(B1, 'A') and Propername(B2, 'N'). The SNePS representation of the *de dicto* rule is shown in Figure 12.6. In our informal predicate notation, it is:

(F6) $$\text{Know}(\texttt{B1}, \text{F}(\texttt{B2})) \ \& \ \text{Know}(\texttt{B1}, \text{Propername}(\texttt{B2}, `N')) \rightarrow \text{F}(\texttt{B2}) \ \& \ \text{Propername}(\texttt{B2}, `N')$$

where Propername(B1, 'A').

It is important to note that, in the presence of (KTB.dr) (i.e., Figure 12.5), (KTB.dd) (i.e., Figure 12.6) is redundant: Our analysis of *de dicto* reports is essentially a conjunction of the representations of two, linked, *de re* reports. Hence, by (RI), two applications of (KTB.dr)—to nodes M14 and M16 of Figure 12.6—yield both consequents of (KTB.dd). In other words, (F6) is a conjunction of two instances of the general form of (F5):

(F5.G) $$\forall \text{P} \forall m \forall n[\text{Know}(m, \text{P}(n)) \rightarrow \text{P}(n)]$$

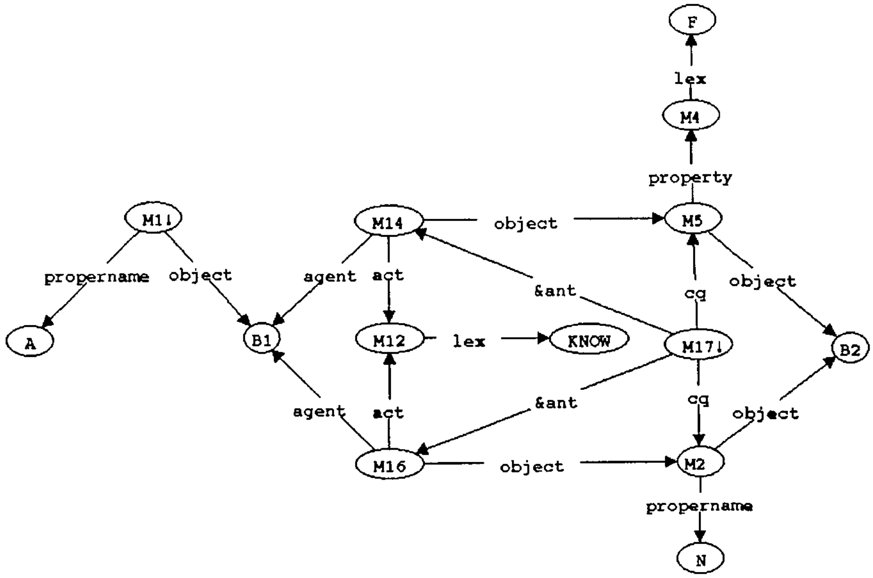

Figure 12.6: (*A* knows that *N* is F) implies *N* is F:
M16: B1 knows M2;
M17!: (Cassie's belief that) (M14 and M16) implies (M2 and M5).

We repeat, for emphasis, that *m* and *n* here do not range over *names* of individuals, but over *individuals*, who may or may not be named or otherwise described.

12.6 *De Se* Knowledge.

However, the veridicality thesis does *not* hold when the objective contains a quasi-indexical (Castañeda 1966: 155; 1967: 93). This can be seen in the general case (from now on, we use '*' instead of the more awkward 'she* (or he*)'):

(V.*) (*A* knows that * is F) → (* is F)

cannot be true, since the occurrence of the quasi-indexical '*' in the *consequent* is not within the scope of an intentional verb; hence, it has no antecedent: We cannot simply detach the consequent, since it cannot stand by itself, so to speak.

William Seager (1990) presents a logic of quasi-indexicals that allows them to occur without antecedents. We think this is incorrect. However, as we have seen, there *are* cases in natural language, in particular, subjective contexts in narrative, in which quasi-indexicals occur without *apparent* antecedent. We maintain, however, that the "subjective agent" of that context serves as its antecedent. (Cf. Banfield 1982; Wiebe & Rapaport 1988; Wiebe 1990ab, 1991, 1994; Fludernik 1993. Recall, too, our discussion of Lespérance and Levesque 1995, in §12.2.2, above.)

That we cannot detach the consequent is even easier to see if we bring Cassie into the picture. In the case of a *de dicto/de se* knowledge report—which, because it is *de dicto*, involves a quasi-indexical—we have:

(KTB.dd-ds)
(Cassie believes that *A* knows that * is F)
→ (Cassie believes that *A* believes that * is F)
& (Cassie believes that she* is F).

(Recall that '*' abbreviates the awkward 'she* (or he*)'. Since we don't know the gender of *A*, we use '*' in the antecedent and first conjunct of the consequent, but since we *do* know Cassie's gender, we can use 'she*' in the second conjunct.) The SNePS representation of part of this rule is shown in Figure 12.7. Informally, it is:

(F7) Know(B1, F(B3)) & Know(B1, Ego(B3)) → F(B3) & Ego(B3).

where Propername(B1, '*A*'). Using (RI), the SNePS Inference Package will assert the propositions labeled M8 and M10 (that is, the consequents of (F7)), thus representing—incorrectly—that Cassie believes that *she** is F. Note, again, that (KTB.dd-ds) is redundant: Using (RI), two applications of (KTB.dr)—to nodes M18 and M19 of Figure 12.7—yield both consequents of (KTB.dd-ds).

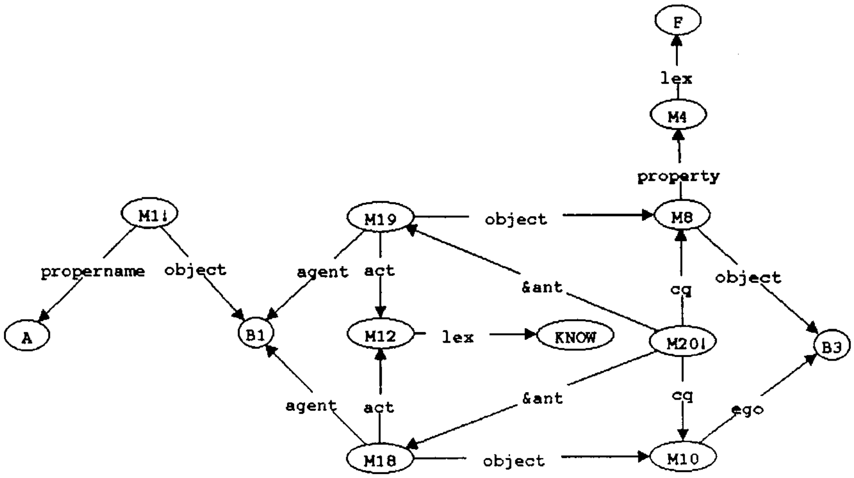

Figure 12.7: (*A* knows that * is F) implies (* is F):
M18: B1 knows M10;
M19: B1 knows M8;
M20!: (Cassie's belief that) (M18 and M19) implies (M8 and M10).

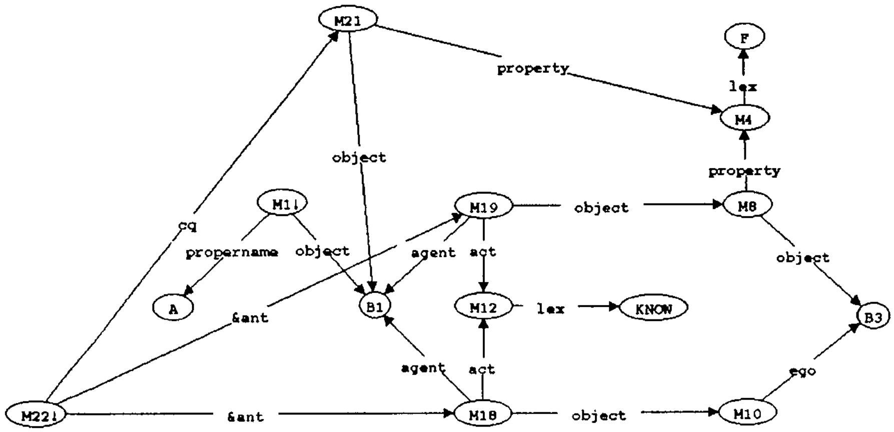

Figure 12.8: (*A* knows that * is F) implies (*A* is F):
`M21`: `B1` is F;
`M22!`: (Cassie's belief that) (`M18` and `M19`) implies `M21`.

Clearly, what we would like is not (KTB.dd-ds), but

(KTB.dd-ds.1)
(Cassie believes that *A* knows that * is F)
→ (Cassie believes that *A* believes that * is F) & (Cassie believes that *A* is F).

part of which can be represented in SNePS as in Figure 12.8. Informally,

(F8) $\quad$ Know(`B1`, F(`B3`)) & Know(`B1`, Ego(`B3`)) → F(`B1`).

where Propername(`B1`, '*A*').

To emphasize that this is the only troubling case, consider a *de re/de se* knowledge report, which, because it is *de re*, does *not* involve a quasi-indexical (Rapaport 1986: 406). We have:

(KTB.drds)
(Cassie believes that *A* knows of her/himself that F)
→ (Cassie believes that *A* believes of her/himself that F)
& (Cassie believes that *A* is F).

In this non–quasi-indexical, *de re/de se* case, we have the same consequent as in the non–quasi-indexical, non-*de se*, *de re* case (KTB.dr): For the antecedent of (KTB.drds)

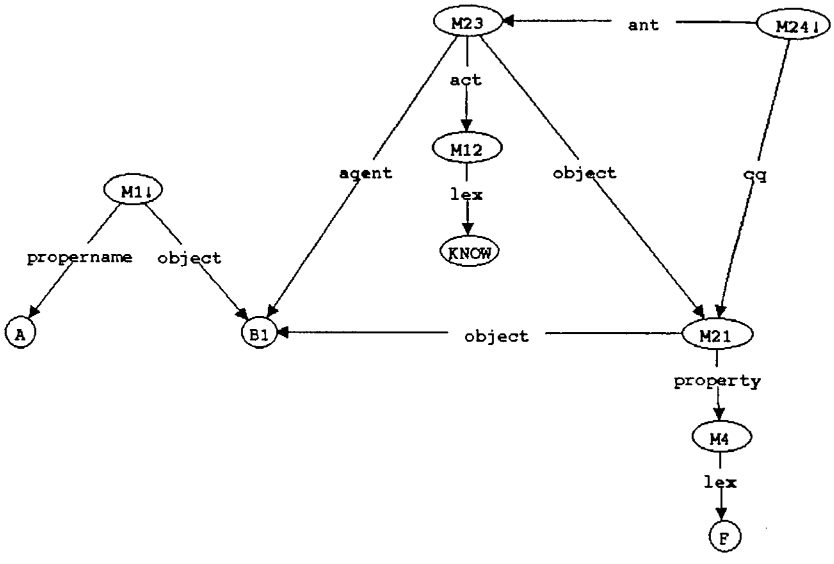

Figure 12.9: (*A* knows of *A* that F) implies (*A* is F):
M23: B1 knows M21;
M24!: (Cassie's belief that) M23 implies M21.

is equivalent (by referential transparency) to: Cassie believes that *A* knows of *A* that F. (See Figure 12.9; informally: Know(B1, F(B1)) $\rightarrow$ F(B1), where Propername(B1, '*A*').) But the consequent of (KTB.dd-ds)—the quasi-indexical, *de dicto/de se* case—is *not* the same as in the non-quasi-indexical, *de dicto* case (KTB.dd). In the former, Cassie believes that *A* believes that * is F; in the latter, Cassie believes that *A* believes that *A* is F.

The main problem is this: It will not suffice to have a separate rule, namely (KTB.dd-ds.1), for the quasi-indexical case, since the rule for the *de re* case (KTB.dr)—and hence the rule for the *de dicto* case, (KTB.dd)—will still allow the inference that we don't want. That is, (KTB.dd-ds)—which is what we don't want—is just a special case of (KTB.dd) and, hence, of (KTB.dr).

12.7 A Solution.

The broader context of our problem is this: In our earlier work (Rapaport & Shapiro 1984, Rapaport 1986), we argued that quasi-indexical reference must be capable of being handled by a belief-representation system, and we presented a computationally adequate mechanism for doing this. That mechanism was adequate as long as we only considered belief reports *in isolation.* When we turn to belief reports that are embedded in surrounding discourse, as in narrative text, the data become more complex, and a correspondingly more complex theory is needed. This is especially the case where conjunctions—especially *sequences*—of belief reports are considered. In Wiebe & Rapaport 1986, we showed that with such sequences, the notions of referential and propositional opacity and transparency interact in ways that blur the distinctions among them. In this paper, we show that our original representation of quasi-indexicals must be modified in order to handle knowledge reports, which are, in fact, conjunctions of belief reports.

The solution we now propose is to represent quasi-indexical, *de se/de dicto* belief and knowledge reports as shown in Figure 12.10. Informally,

(I10.B) Propername(B1, '*A*') & Believe(B1, F(B1)).

(I10.K) Propername(B1, '*A*') & Know(B1, F(B1)).

(I10.B) *replaces* (I3). Notice that there is no "Ego belief" component as in (I3). Using the representation of (I10.K), the inference from

Cassie believes that *A* knows that * is F

to

Cassie believes that *A* is F

can be handled by the *same* rule (KTB.dr) as in the other cases (roughly because 'Propername(B1, '*A*')' is outside the scope of 'Know'), and—because there is no "Ego belief" component—'Cassie believes that she* is F' is no longer inferable.

However, there are several potential problems that must be cleared up before this solution can be adopted. First, Figure 12.10 is a representation for quasi-indexicals that was *rejected* in our earlier work! So, we must re-examine those arguments. Second, our original representation made use of an ego arc and a representation of *A*'s "self-concept", whereas our new representation does not. But the notion of an agent's self-concept is of independent importance, so we must explore alternative representations for it. Third, the representation in Figure 12.10 does not appear to be *de dicto* (since it does not consist of *two*, linked, *de re* belief reports); so we must re-examine the nature of *de dicto* belief reports to see whether our claim that quasi-indexical belief is *de dicto* can be maintained. We now turn to an exploration of these issues.

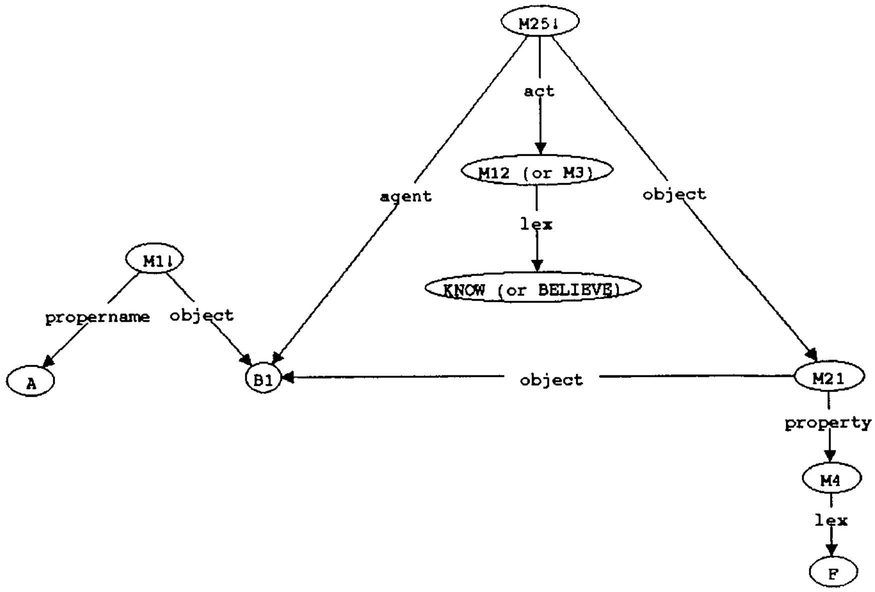

Figure 12.10: *A* knows (or believes) that * is F:
M25!: (Cassie's belief that) B1 knows (or believes) M21.

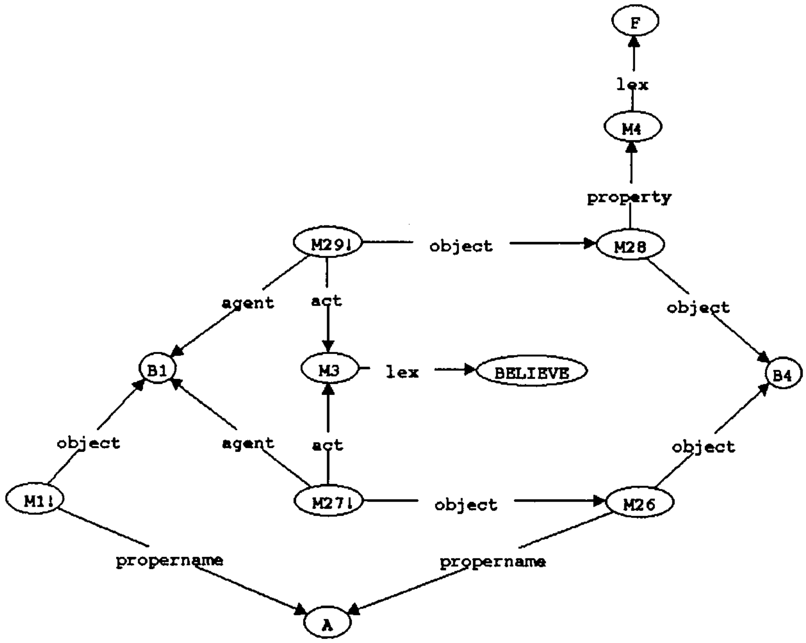

Figure 12.11: *A* believes that *A* is F:
M26: B4's name is '*A*';
M27!: (Cassie's belief that) B1 believes M26;
M28: B4 is F;
M29!: (Cassie's belief that) B1 believes M28.

12.7.1 Is Figure 12.10 Acceptable?

In Rapaport & Shapiro 1984 and Rapaport 1986, we rejected the representation of Figure 12.10 on two grounds. The first was that it ambiguously represented both

(7) *A* believes that * is F

and

A believes that *A* is F,

which are not equivalent. But the latter really should be represented as in Figure 12.11, informally as:

(F11) Propername(`B1`, '*A*') & Believe(`B1`, Propername(`B4`, '*A*'))
& Believe(`B1`, F(`B4`)).

So, the representation of Figure 12.10 *is* available to represent the former. The Figure-12.10 representation is ambiguous only if `B1` is interpreted as a name, which we do not do. This issue is taken up in §12.7.2.

We also argued that the Figure-12.10 network did not adequately represent the quasi-indexical nature of the belief report, on the grounds that node `B1`—representing *A*'s self-concept—was both inside and outside the intentional context—that is, in both Cassie's and *A*'s belief spaces. But, of course, *all* nodes are in Cassie's belief space, and what must be represented is *Cassie's* belief, which is that the person believed by *A* to be F is *A*—the believer—herself or himself. Figure 12.10 does represent this. What it does not—and should not—do is suppose that *A* characterizes her- or himself with the *name* '*A*'.

12.7.2 The Proper Treatment of Proper Names.

The original motivation, however, for the Figure-12.3 representation was *not* the alleged ambiguity of Figure 12.10, but the *actual* ambiguity of Maida and Shapiro's representation (1982), shown in Figure 12.12. Here, it should be noted, the `propername-object` case frame is not used. Informally,

(F12) Believe(*A*, F(*A*)).

Note that here '*A*' is a constant; it is *not* a proper name.

The proper lesson to be learned from this is the importance of the `propername-object` proposition for the representation of cognitive agents. Shapiro used such propositions before the Maida and Shapiro paper (using a `name-named` case frame; Shapiro 1975, 1979a, 1982), but felt that nothing major was lost by abbreviating the representation used in Maida & Shapiro 1982 to the extent of not separately showing this proposition. It was the abbreviated version that Rapaport realized was ambiguous between the *de re* and *de dicto* cases, and this led us to the Ego

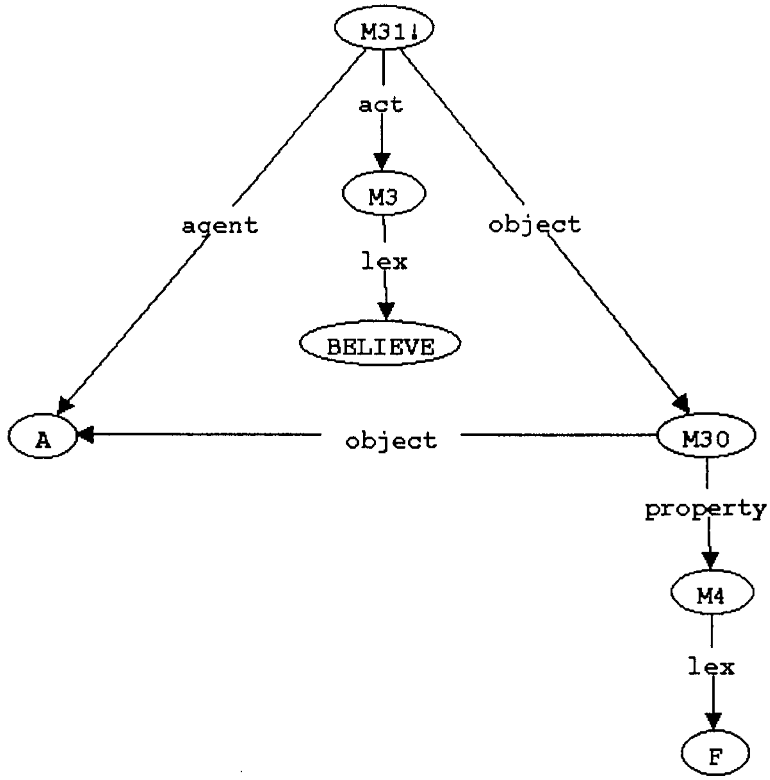

Figure 12.12: A Maida & Shapiro network for '*A* believes that * is F':
M30: *A* is F;
M31!: (Cassie's belief that) *A* believes M30.

proposition (as reported in Rapaport & Shapiro 1984 and Rapaport 1986). We can now see that, although the Ego proposition works when representing nested beliefs, it does not work when representing nested knowledge.

The lesson is this (as Wilks et al. 1989: 499 and Ballim & Wilks 1991: 77–78 observe): When representing a cognitive agent within a belief system, it is important to represent the agent in a way that is neutral to any properties (including its name) ascribed to it by the believer. In that way, the representation of the cognitive agent may be used in representations of its beliefs about itself without automatically ascribing to it any of the properties ascribed to it by the believer. If the representation is not neutral, and the automatic transfer of the property ascription is not wanted, node splitting must be used (see Maida & Shapiro 1982).

Adam J. Grove (1995) also argues that it is important to distinguish between an agent and a name or description for the agent: "It is not enough to know *who* a name refers to— ... we must also decide *how* the reference is made" (p. 314). Although he is not concerned (as we are) with natural-language understanding or cognitive modeling (cf. p. 320), the examples he cites are similar to those cited in Rapaport 1986, and he also notes the importance of quasi-indexicals, though without calling them by that name: "an individual's way of referring to *itself* seems to have special properties" (p. 326; cf. p. 318). He introduces "a special name *I* that allows the agent to refer to himself" (sic; p. 319), which is similar to our notion of an " 'I'-pointer" (see §12.7.3, below), and he introduces a "special symbol" **me**, which plays a role similar to Lespérance and Levesque's **self**.

Both *I* and **me** have quasi-indexical features: "The best reading of our *I* depends on context; for instance, we would read $K_{\mathbf{n}}K_{\mathbf{m}}K_{I}\varphi$ as '**n** knows that **m** knows that *he himself* knows φ' " (p. 319; italics in original), and

> ... **me** ... denotes the agent *a* from whose viewpoint [possible world] *w* is being considered, and so functions very much like ... "*I*" .... The difference between *I* and **me** is minor: the former is a name that usually denotes the identity relation while the latter is of sort agent. In practice, the two can be regarded similarly. (p. 328.)

In the formal development of his system, Grove has the following axiom (p. 335):

(M2) $$K_t\varphi \Rightarrow \varphi[t/\mathbf{me}] \text{ if } t \text{ is substitutable for } \mathbf{me},$$

where "if t and t' are terms, by $\varphi[t/t']$ ... we mean a formula like φ, except that all ... 'substitutable' occurrences of t are replaced by t' " (p. 335). This is an error (which we have confirmed with Grove (personal communication, 3 September 1995)): The substitution notation in this passage should have been: $\varphi[t'/t]$. The point is that an occurrence of **me** in φ in the scope of K_t means t, i.e., "he himself": (M2) says that if t knows that he himself (or she herself) satisfies φ, then t satisfies φ.

However, although Grove's theory may solve the problem that Lespérance and Levesque have, (M2) puts indexicals in the formal representation language. Hence, Grove has a non-compositional semantics, since **me** refers to different things in different contexts. We think that our solution is better: All terms have fixed denotation, and proposition φ is the *same* φ in both antecedent and consequent of our version of (M2) (viz., (KTB.dr)). The symbol that represents the φ that Cassie believes that t knows is the same symbol that Cassie believes, not a copy of it. Grove's is not only a copy, but a changed copy, t/**me**. So, assuming a Lisp representation, our φs are "eq", but Grove's is not even "equal". In Grove's system, if we ask, "Do you believe the *same* thing that t believes?", the answer is: not the *same*, but *similar*. In ours, by contrast, it is the *same*.

12.7.3 How to Represent a Self-Concept.

With the ego arc, we were able to represent Cassie's beliefs about herself. It is essential that we be able to do this. Not only must we be able to represent Cassie's belief, say, 'I am intelligent', but Cassie might have false nested beliefs about herself or fail to believe that she in fact has certain beliefs about herself. For example, Cassie might explicitly believe that she believes that P, yet she might not in fact believe that P (as evidenced by her failure to act in accordance with P). Or Cassie might in fact believe that P, yet not believe that she believes it (or, of course, believe that she does not believe it).

How can we represent these without the ego arc? The solution we have chosen is a generalization of a mechanism that our research group uses for representing the temporal indexical 'now': A node representing 'now' is identified by a (movable) 'now'-pointer. The "temporal" node pointed to by the 'now'-pointer changes as linguistic cues in the discourse or narrative move the 'now' point along (cf. Almeida & Shapiro 1983; Almeida 1987, 1995; Shapiro & Rapaport 1987).

Similarly, within Cassie's belief space, we postulate an 'I'-pointer, which, at the beginning of a dialogue with Cassie, is initialized to point to a node, which will then represent Cassie's self-concept. Unlike the 'now'-pointer, the 'I'-pointer does not need to be updated. On the other hand, just as, when reading a narrative, 'now'-points are stacked when entering sub-narratives (for example, a flashback), the 'I'-pointer is stacked when entering nested belief spaces. At the top level, the word 'I' is used to express the node pointed to by the 'I'-pointer; when the context is a nested belief space, the word 'I' would change to 'she*' or 'he*' (as appropriate).

Haas, in his sentential theory of indexical expressions (1993), has also noted the importance of quasi-indexicals. Instead of a self-*concept*, his analysis requires a "self-name" for each agent. This is "a standard constant" used by the speaker "to refer to himself or herself" (p. 646). Thus, presumably, an agent John's own representation of his utterance "I am smart" would be something like:

smart(i)

or, perhaps

believe(i, smart(i))

where 'i' is John's selfname. A selfname, of course, is not a quasi-indexical, since a quasi-indexical is an expression used by *another* speaker to represent someone else's selfname. To represent sentences containing quasi-indexicals, Haas requires quite a bit of formal machinery:

> An utterance of "John thinks he is smart" would normally express a singular proposition (Q, f), where Q is the wff
>
> 23 denote(z, john) $\wedge$ believe(john, subst('smart(x), ['x], [z]))
>
> and f is a function that maps the variable z to John's selfname. (p. 646.)

Here, subst('smart(x), ['x], [z])) is the wff that results from simultaneous substitution of term z for all free occurrences of variable 'x in 'smart(x); i.e., it is 'smart(z).

How does this compare with our treatment of quasi-indexicals? At first glance, the 'denote' predicate seems to play the same role as our `object-propername` caseframe. However, z is not the object whose proper name is 'John'; rather, z represents John's selfname, and, in fact, Haas's analysis of this utterance seems to interpret 'john' as John himself and not a name for John, as both our theory and Grove's would have it. That is, Haas's analysis of "John thinks he* is smart" does not represent the fact that the believer is *named* 'John'. So, z seems to be the utterer's representation of John's selfname, something our theory avoids the need for in this context. (If needed, however, we could easily represent propositions such as that John believes that his* name is 'N', or that 'i' is his selfname.) As for the second conjunct of (23), Haas's sentential belief-representation theory tries to represent the (mental?) *sentence* that John believes, hence the need for the substitution mechanism. We have presented arguments elsewhere (Shapiro 1993) for the advantages of propositional representations over sentential ones.

12.7.4 Is Figure 12.10 *De Dicto*?

Quasi-indexical *de se* belief reports are *de dicto*. This is, perhaps, arguable (cf., e.g., Castañeda 1989: 16, n.10). But, like *de dicto* and unlike *de re* belief reports, they are referentially opaque and propositionally transparent, at least in isolation. Yet Figure 12.10 does not have the structure of a representation of a *de dicto* report; indeed, it appears to have the structure of the representation of a (single) *de re* report.

Now, the *de dicto/de se* report

(7) *A* believes that * is F

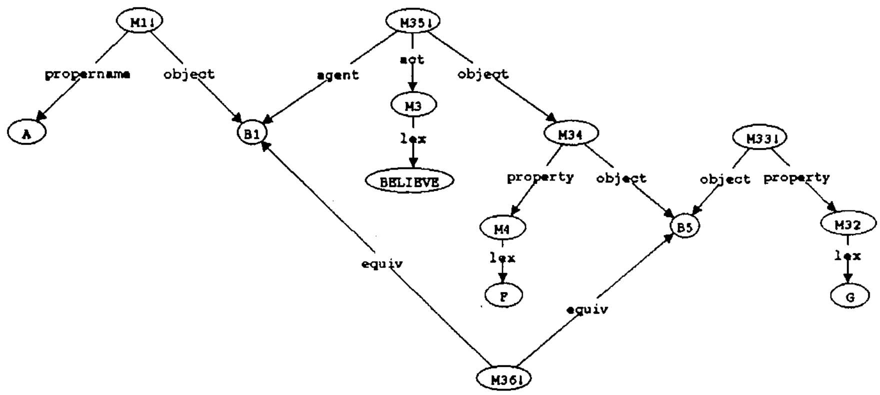

Figure 12.13: Cassie's belief that: *A* believes of someone whom Cassie believes to be *A* (viz., someone whom Cassie believes to have property G) that F:
`M33!`: (Cassie's belief that) `B5` is G;
`M34`: `B5` is F;
`M35!`: (Cassie's belief that) `B1` believes `M34`;
`M36!`: (Cassie's belief that) `B1` $\equiv$ `B5`.

implies, but is not implied by, the *de re/de se* report

(12.10) *A* believes of her/himself that F.

Figure 12.10 is the representation of (12.7); it is also *a* representation of (12.10), which is consistent with the fact that (12.7) implies (12.10). But in various contexts, various representations will be used to represent (12.10) (besides Figure 12.10, for example, there could be Figure 12.13). So it is not the case *in general* that (12.10) implies (12.7).

12.7.4.1 Castañeda-style predication.

Is there, though, a way to represent the quasi-indexical *de se* belief in such a way that it wears its *de-dicto*–ness on its sleeve, so to speak? There is, but it might not serve any purpose. Our (F2)-analysis of *de dicto* belief reports is this:

(5) *A* believes that *N* is F

is analyzed as (a Skolemized form of):

A believes that something that is named '*N*' is (the same as something that is) F.

Similarly, our (F3)-analysis of

(7) *A* believes that * is F

is (a Skolemized form of):

A believes that something that is * is (the same as something that is) F.

These suggest the patently *de dicto* SNePS networks of Figures 12.14 and 12.15. The mode of predication exhibited here is not a simple `object-property` case frame. Rather, '*N* is F' is analyzed as (a Skolemized form of):

(C) $\exists x \exists y[x$ is named 'N' & y is F & x equiv $y]$

This is very close to the theory of predication put forth in Castañeda 1972 [1974] (cf. 1989) (where the Skolem constants would now be interpreted as designating "guises"—Castañeda's version of intensional entities), and which we previously suggested as an analysis of predication in SNePS (Rapaport 1985). It can now be seen to have the additional advantage of exhibiting the *de dicto* nature of quasi-indexical *de se* reports.

(There is another alternative, viz., a Skolemized form of:

$\exists x[x$ is named 'N' & x is F$]$.

This is not exactly in the spirit of that aspect of Castañeda's theory according to which each guise has exactly one identifying feature, with various equivalence relations ("consubstantiation", "consociation", etc.) linking guises together into complex, lattice-like structures. (For details, see Castañeda 1972 [1974], 1989; Tomberlin 1983, 1986.) Nevertheless, this alternative satisfies the reasonable principles that Cassie can identify whom Fness is being predicated of and that to say something about an entity, one needs two ways to characterize it: one to identify it and one to predicate something of it.)

Does (C) run into the same problem that our earlier *de dicto/de se* representation does with respect to knowledge? No: 'Cassie believes that *A* knows that * is F' would simply imply that Cassie believes that someone is F and that that someone is *A*, which is precisely right.

So, is the extra belief about the equivalence of the object that is F and the object named '*N*' needed? If not, then the representation of Figure 12.10 suffices (at least until more complex data are unearthed). We think that it is *not* needed, at least in order to render the Figure-12.10 analysis *de dicto*. But to show this, we advocate a new understanding of *de dicto* and *de re* belief reports in the context of conversation.

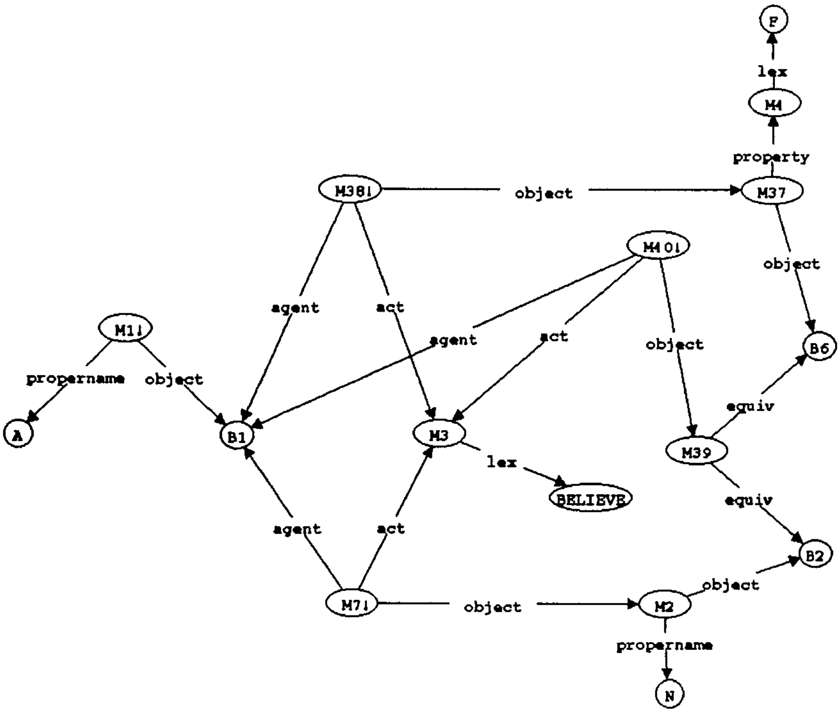

Figure 12.14: *A* believes (*de dicto*) that something named '*N*' is (the same as something that is) F:
M37: B6 is F;
M38!: (Cassie's belief that) B1 believes M37;
M39: B2 ≡ B6;
M40!: (Cassie's belief that) B1 believes M39.

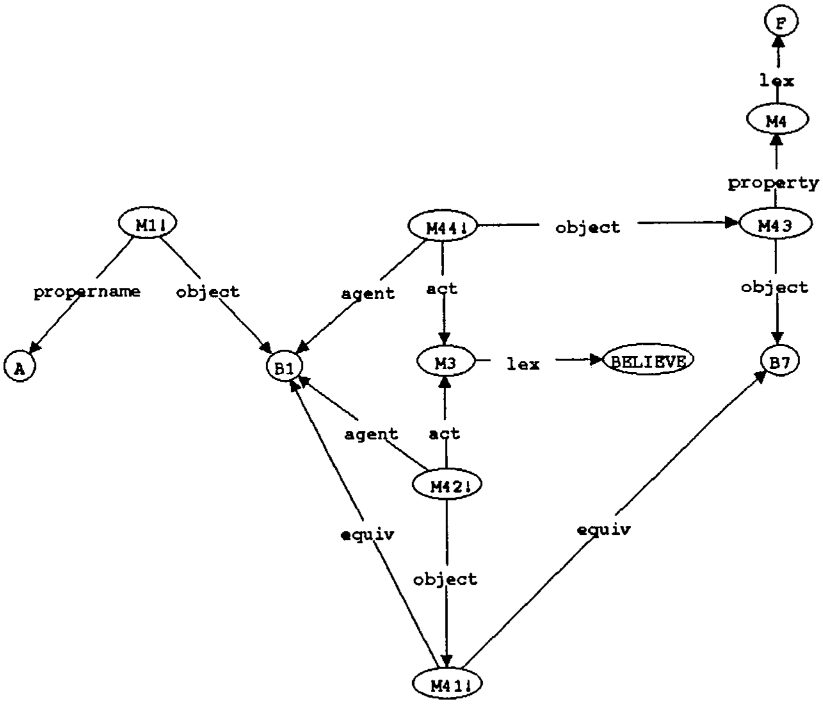

Figure 12.15: *A* believes (*de dicto*) that something (viz., B1) that is * is (the same as something (viz., B7) that is) F:
M41!: (Cassie's belief that) B1 ≡ B7;
M42!: (Cassie's belief that) B1 believes M41;
M43: B7 is F;
M44!: (Cassie's belief that) B1 believes M43.

12.7.4.2 A new theory of *de re* and *de dicto* belief reports.

Consider two participants in a dialogue, Cassie and Oscar. Suppose that Oscar says to Cassie (perhaps in a vain attempt to impress her),

I am rich,

thus expressing the belief represented in Figure 12.16A. Cassie's *interpretation* of this is expressed by her as

Oscar believes that he* is rich

and represented (using the Figure-12.10 representation) as in Figure 12.16B. Suppose, next, that Oscar says to Cassie (perhaps in a vain attempt to make her jealous),

Lucy is sweet,

thus expressing the belief represented in Figure 12.17. Cassie's *interpretation* of this is expressed by her as

Oscar believes that Lucy is sweet

and represented as in Figure 12.18.

That is, representations of *de dicto* belief reports are *Cassie's interpretations* of reports made *by* the believer (that is, reports *from* the believer *to* Cassie (including reports *about* her- or himself)), and are such that Cassie's representation is "exactly" like the believer's representation, except for two facts: (1) In the representation, all nodes are in Cassie's belief space, not the believer's. (2) In the verbal expression, the deictic shift from *indexicals* (used by the believer) to *quasi*-indexicals (used by Cassie) only occurs in Cassie's attempt to express her understanding of Oscar's belief (or a third person's attempt to express to Cassie his understanding of Oscar's belief). The analogue in the representation of this shift in the verbal expression is the use of an embedding belief-structure (i.e., "Oscar believes that ...") in Cassie's belief space in place of the 'I'-pointer in Oscar's belief space. That is, Oscar's 'I'-point (i.e., his self-concept node) maps into Cassie's node representing Oscar (an 'Oscar-point', so to speak).

Finally, suppose that a third cognitive agent, Boris,[2] knows that the person Oscar believes to be Mitzi is the person Cassie believes to be Mary (represented by node `M211` in Figure 12.20). That is, Boris believes that Oscar believes *that* Mitzi is tall, and Boris believes that Oscar believes *of* Mary that she is tall, represented in Figures 12.19 and 12.20. Suppose that Boris tells Cassie that Oscar believes *of* Mary that she is tall. Cassie's *interpretation* of Boris's third-person, *de re* report is that Oscar believes *of* Mary that she is tall, represented in Figure 12.21. That is, a *de re* belief report is *Cassie's interpretation* of a *third* person's interpretation of Oscar's beliefs (that is, a

[2]The *B*orrowed *I*ntelligent *S*ystem (name borrowed from, and with apologies to, Lehnert et al. 1983).

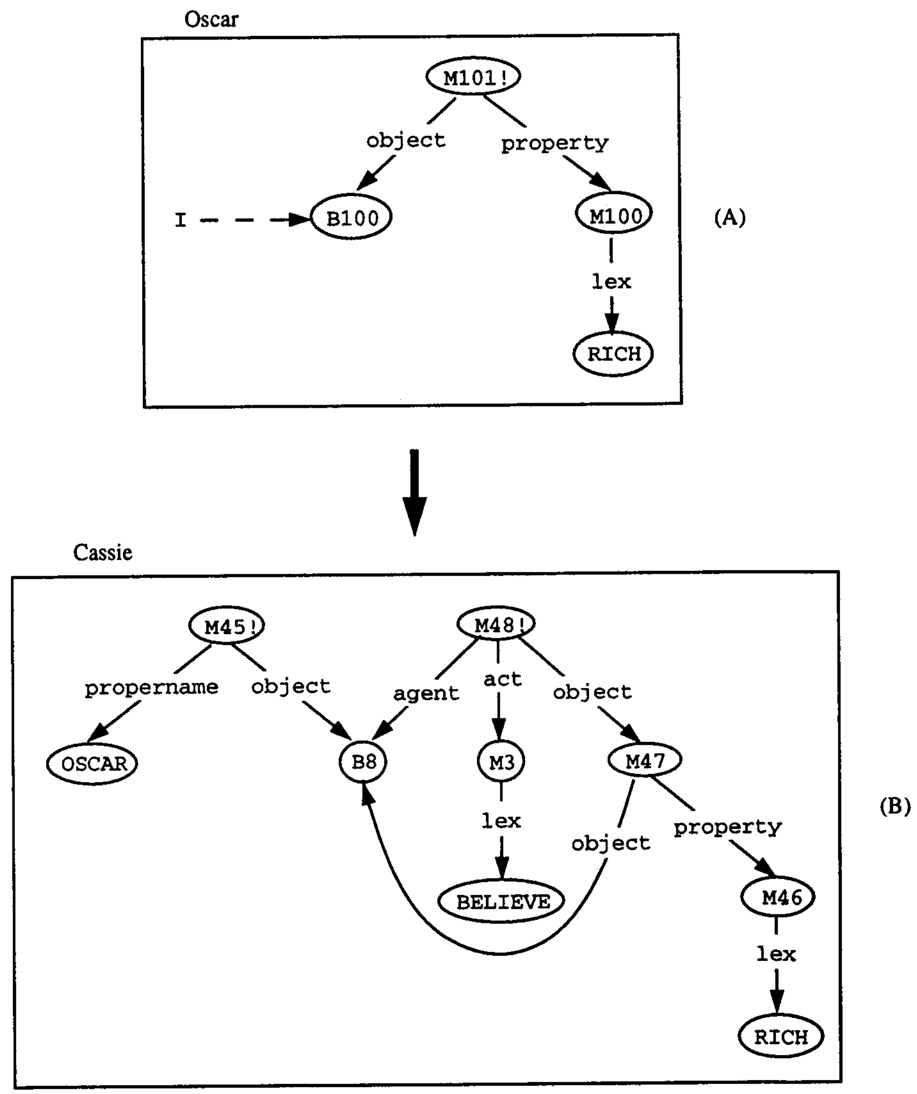

Figure 12.16: (A) Oscar's belief: I am rich.
`M101!`: (Oscar's belief) I am rich (i.e., Oscar's belief that he* is rich).
(B) Cassie's belief: Oscar believes that * is rich.
`M45!`: (Cassie's belief that) `B8`'s name is 'Oscar';
`M47`: `B8` is rich;
`M48!`: (Cassie's belief that) `B8` believes `M47`.

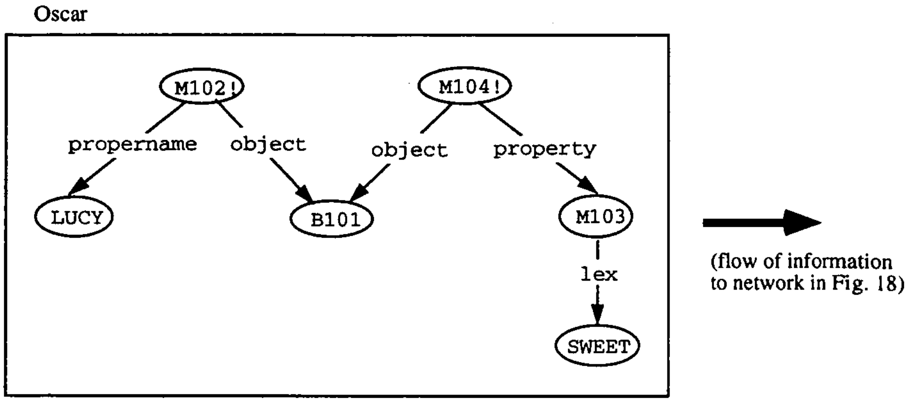

Figure 12.17: Oscar's belief: Lucy is sweet:
`M102!`: (Oscar's belief that) `B101`'s name is 'Lucy';
`M104!`: (Oscar's belief that) `B101` is sweet.

report *from* a third person *to* Cassie *about* the believer), and is such that Cassie's representation is like Oscar's only with respect to the fragment that is in common. This is the core of what is meant by 'propositional opacity'. (*De re* reports might also be *inferred* by Cassie from other beliefs that she has.) (The reader is reminded of our convention about the use of 'of' and 'that' in belief reports; see §12.2.1. By the use of this convention, there is no ambiguity, despite the fact that, *in ordinary conversation*, Cassie will understand 'Mary' as reflecting Boris's beliefs (cf. Fodor 1970/1979, Wiebe 1991) and faces the ambiguity: Does Boris believe that Oscar believes she is named Mary? If yes, this is what is called "*de dicto* in conversation" in Wiebe 1991; if not, it is "*de re* in conversation".)

There is one final issue to consider. Suppose that Cassie is told by Boris that Oscar believes of the person whom Cassie and Boris believe is Oscar that he is rich. Should Cassie interpret this as in Figure 12.22 or Figure 12.23 (cf. Figures 12.10, 12.13, respectively)? If Cassie interprets Boris's belief report as in Figure 12.22, then she could infer that Oscar believes that he* is rich, which might be false. So Figure 12.23 ought to be Cassie's interpretation. If Boris then tells Cassie that Oscar believes that he* is rich, then Figure 12.23 would be modified as in Figure 12.24, because Cassie still does not know whether Oscar believes that two people or one person is rich. Finally, if Boris tells Cassie that Oscar only believes himself* to be rich, then Cassie must "merge" two nodes (`B8` and `B10`) in her representation of Oscar's belief space, as in Figure 12.25 (cf. Maida & Shapiro 1982).

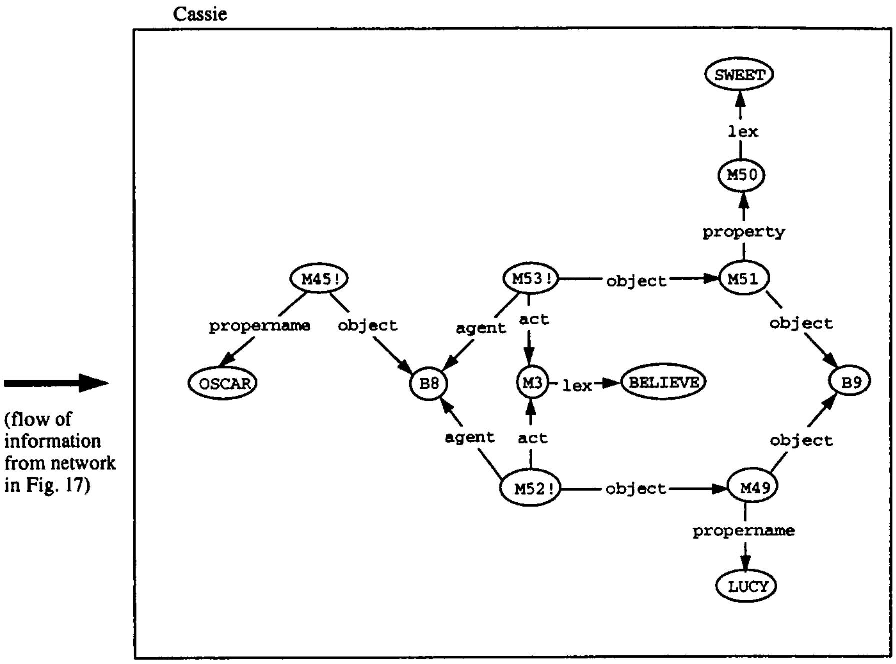

Figure 12.18: Cassie's belief: Oscar believes that Lucy is sweet:
M49: B9's name is 'Lucy';
M51: B9 is sweet;
M52!: (Cassie's belief that) B8 believes M49;
M53!: (Cassie's belief that) B8 believes M51.

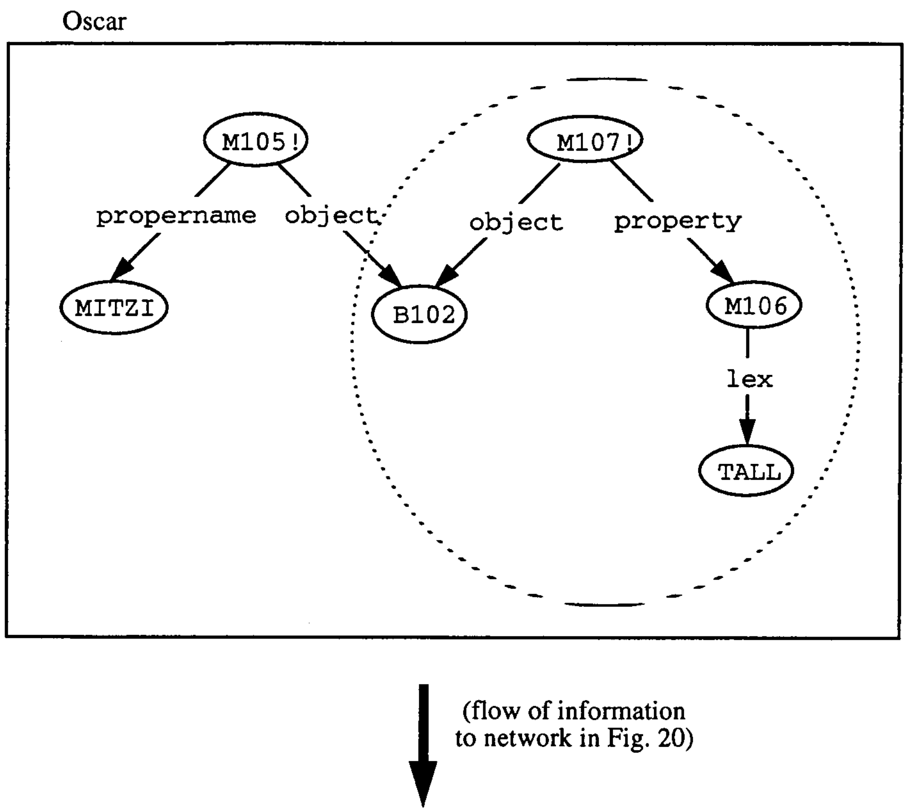

Figure 12.19: Oscar's belief: Mitzi is tall:
`M105!`: (Oscar's belief that) `B102`'s name is 'Mitzi';
`M107!`: (Oscar's belief that) `B102` is tall.

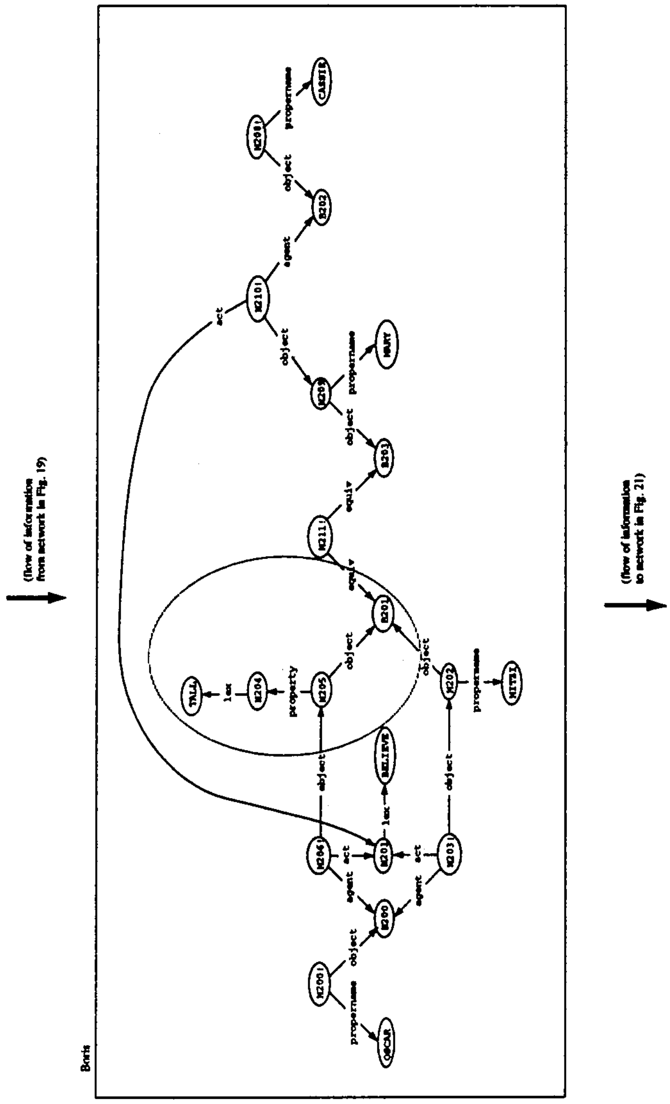

Figure 12.20: Boris's belief: Oscar believes that Mitzi is Tall and Oscar's Mitzi is Cassie's Mary.
M200!: (Boris's belief that) B200's name is 'Oscar';
M202: B201's name is 'Mitzi';
M203!: (Boris's belief that) B200 believes M202;
M205: B201 is tall;
M206!: (Boris's belief that) B200 believes M205;
M208!: (Boris's belief that) B202's name is 'Cassie';
M209: B203's name is 'Mary';
M210!: (Boris's belief that) B202 believes M209;
M211!: (Boris's belief that) B203 ≡ B201.

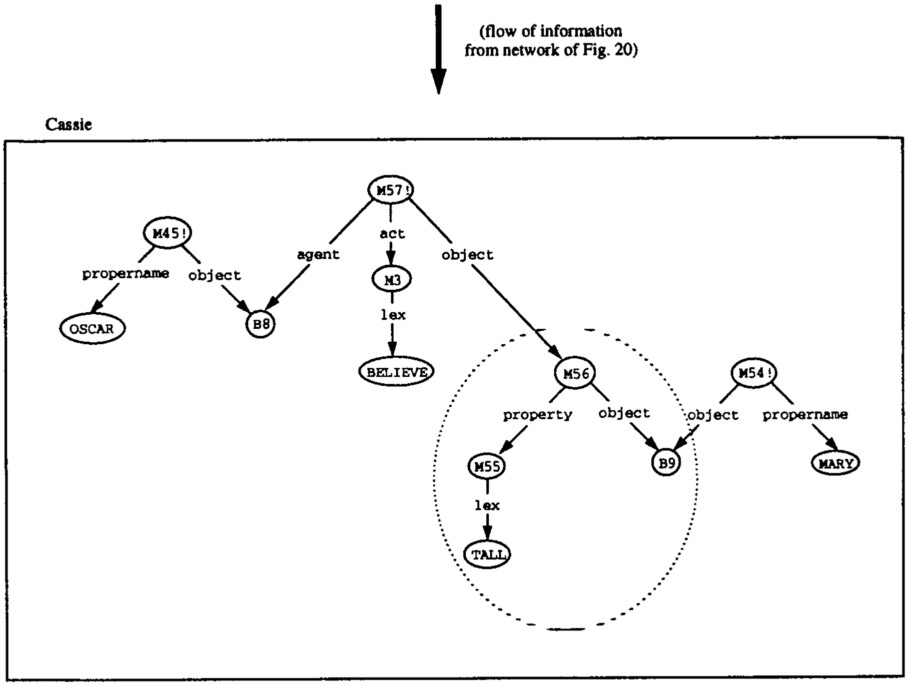

Figure 12.21: Cassie's belief: Oscar believes of Mary that she is tall.
`M54!`: (Cassie's belief that) `B9`'s name is 'Mary';
`M56`: `B9` is tall;
`M57!`: (Cassie's belief that) `B8` believes `M56`.

(N.B.: In Figs. 12.19–12.21, the networks in dashes correspond to each other.)

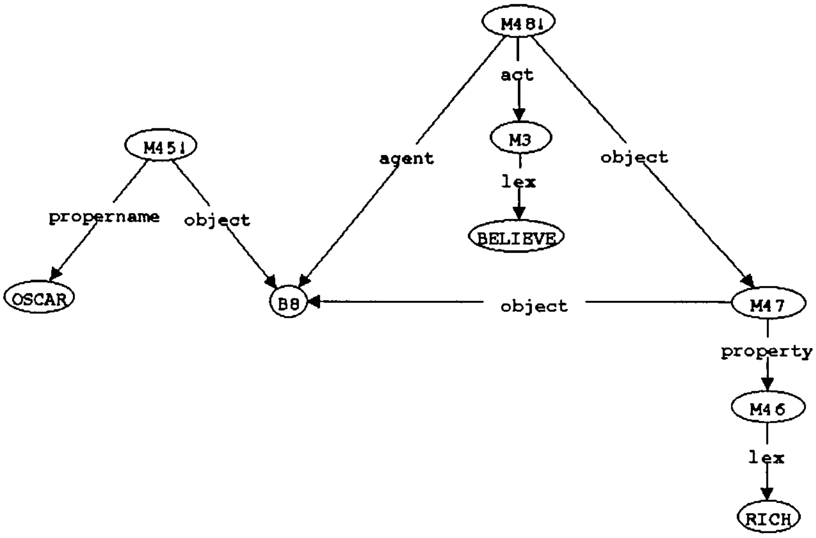

Figure 12.22: Oscar believes of Oscar that he is rich (?)

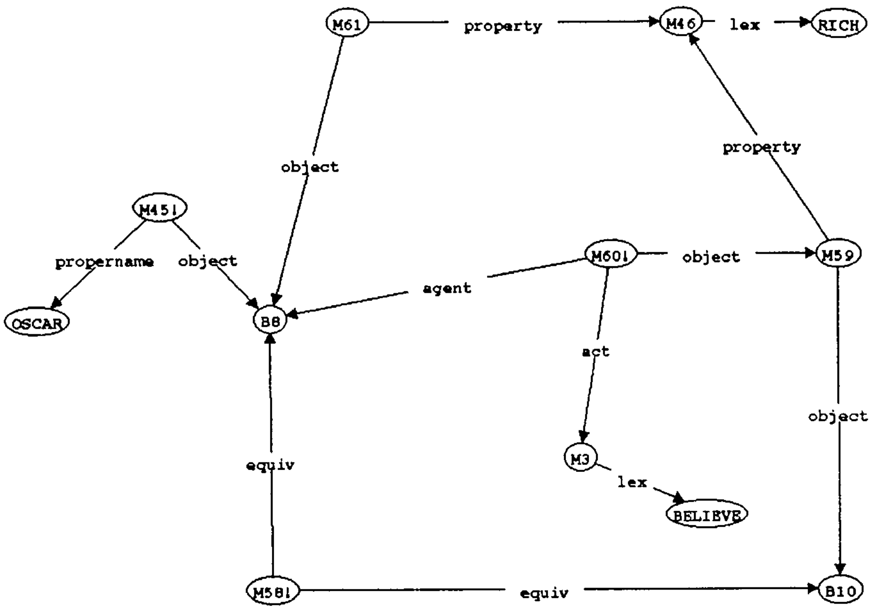

Figure 12.23: Oscar believes of Oscar that he is rich (?)
M59: B10 is rich;
M58!: (Cassie's belief that) B8 $\equiv$ B10;
M60!: (Cassie's belief that) B8 believes M59.

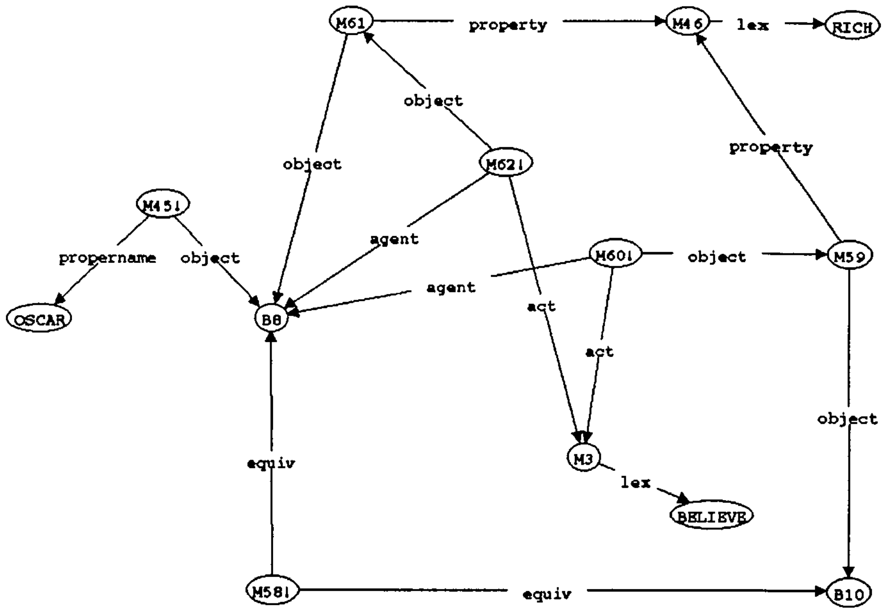

Figure 12.24: Oscar believes of (the person whom Cassie believes is) Oscar that he is rich (M60!)
and Oscar believes that he* is rich (M62!).

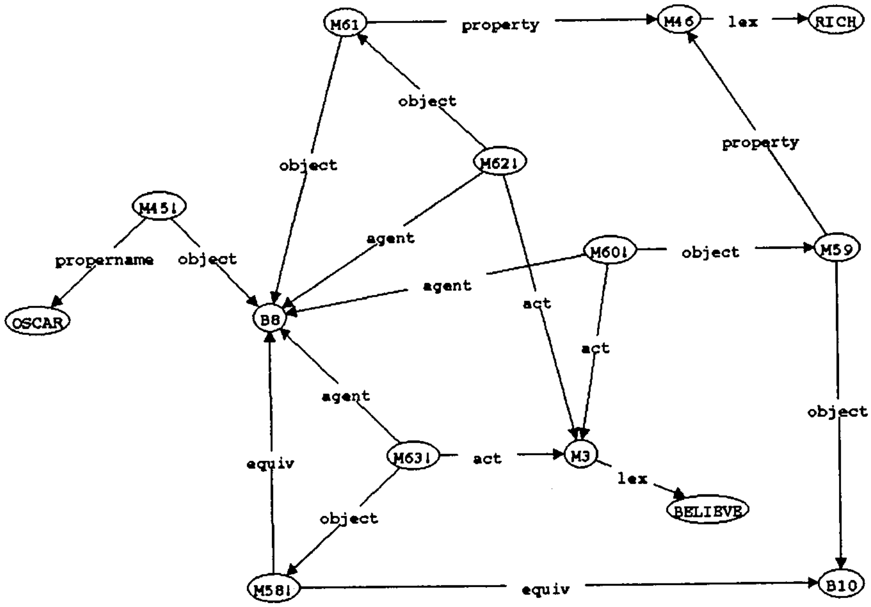

Figure 12.25: Oscar believes of (the person whom Cassie believes is) Oscar that he is rich (M60!)
and Oscar believes that he* is rich (M62!)
and Oscar believes that he* is the person whom he believes to be rich (M63!)

12.8 Replies to Wilks, Ballim, and Dietrich.

The revised representation of *de se* belief and knowledge reports, together with the 'I'-pointer, answers several objections raised by Wilks, Ballim, and Dietrich (1989; cf. Ballim & Wilks 1991: 73–81).

We claimed (in Rapaport 1986) that a knowledge-representation system must correctly represent quasi-indexicals in order to avoid invalid inferences such as those discussed in §12.2, above. Wilks et al. observe that these "are not representation problems, rather they are problems of ENGLISH usage" (Wilks et al. 1989: 499; Ballim & Wilks 1991: 77), that is, problems of *interpreting* and *generating* belief reports. We agree that these are problems of interpretation and generation. In particular, we argued in Rapaport 1988b (cf. Rapaport 1995) that such semantic interpretation of a belief report into a representation in a knowledge base is a syntactic process (as is generation from the language of thought in the knowledge base to an expression in the public communication language (Shapiro 1993)). The belief report (of the speaker) is represented in the knowledge base (of the hearer).

It does not, of course, follow from this necessarily that every *part* of the report is represented. That would probably only be the case if what was represented was the linguistic expression of the report, rather than its propositional content. Thus, if the report contains a quasi-indexical, it does not follow that the linguistic expression that is the quasi-indexical need be represented in a language of thought that captures propositional content. But the *contribution* of the quasi-indexical must somehow be represented. That is, the representation must accurately reflect the meaning of the report, and, insofar as the quasi-indexical contributes to the meaning, the representation must represent that contribution. (Cf. Russell's (1905) analysis of definite descriptions, which represents the contribution of 'the' without representing 'the' *per se*: 'the F is G' is analyzed as "one and only one thing is F, and it is (also) G".) When Wilks et al. argue that quasi-indexicals need not be represented, they mean that there need not be a *node* corresponding to the quasi-indexical, such as the node at the head of an `ego` arc. Although we no longer need such a node, we would still say that we are "representing" quasi-indexicals.

It is perhaps worth recalling in this connection that the fact that quasi-indexicals in English are morphologically indistinguishable from pure indexicals is a poverty of English. If Cassie spoke Tuburi (Sells 1987: 446ff) or Mandarin (Zubin in Rapaport et al. 1989, §4; Li 1991; Li & Zubin 1995), it would be even more obvious that quasi-indexicals would have to be represented, since in those languages, among others, quasi-indexicals are lexically distinct, "logophoric", pronouns.

Wilks et al. are wrong, however, when they say that SNePS networks cannot be partitioned "in the classic sense of" Hendrix 1979 into belief spaces (1989: 500; Ballim & Wilks 1991: 79). First, there is an easy technique to form Hendrix-like partitions. For any set of nodes to be partitioned off, simply build a node with arcs (labeled by a mnemonic for the partition) to each of them. That node then represents the partition. Second, our definition of belief space is quite explicit (cf. Rapaport 1986; Wiebe & Rapaport 1986, 1988; Martins & Shapiro 1988; Wiebe 1990, 1991, 1994; see the examples in §12.5, above). Informally, a belief space of an agent *A* should be the set of nodes that represent all the entities *A* has conceived of. (This is a slightly different use of 'belief space' from that of Shapiro 1993, where the "belief space" of an agent is the set of propositions that agent believes. The current paper is consistent with Shapiro 1993 in having the object of belief or knowledge be a proposition represented by a term of the knowledge-representation language.) But that would simply be *A*'s network, and remember that each agent has his or her own network disjoint from the networks of all other agents. So a more accurate rendition would be that a belief space of an agent *B* within the belief space of an agent *A* is the set of nodes representing all the entities that *A* believes that *B* has conceived of. Nesting further, *A* may believe that *B* believes that *C* has conceived of certain entities. The set of nodes representing those entities would be the belief space of *A*'s *B*'s *C*. Putting this somewhat formally, assuming that we are discussing Cassie's network, and that in Cassie's network k-or-b is either $\{\langle \texttt{lex}, \texttt{believe}\rangle\}$ or $\{\langle \texttt{lex}, \texttt{know}\rangle\}$:

- The entire network forms Cassie's belief space.
- For any agent a, every node n for which

  $$\{\langle \texttt{agent}, \mathsf{a}\rangle, \langle \texttt{act}, \mathsf{k\text{-}or\text{-}b}\rangle, \langle \texttt{object}, \mathsf{n}\rangle\}$$

  is asserted and all nodes dominated by n are in Cassie's a's belief space.
- Every node n for which

  $$\begin{aligned}&\{\langle \texttt{agent}, \mathsf{a}_1\rangle,\\ &\ \langle \texttt{act}, \mathsf{k\text{-}or\text{-}b}\rangle,\\ &\ \langle \texttt{object}, \ldots \{\langle \texttt{agent}, \mathsf{a}_m\rangle, \langle \texttt{act}, \mathsf{k\text{-}or\text{-}b}\rangle, \langle \texttt{object}, \mathsf{n}\rangle\} \ldots\rangle\}\end{aligned}$$

  is asserted and all nodes dominated by n are in Cassie's a_1's ... a_m's belief space.

Thus, when representing reports about another cognitive agent's beliefs, it is important to remember that we are modeling *Cassie's* mind, not two separate minds (Cassie's and Oscar's): Thus, it is a feature—not a bug—that Cassie's representation of Oscar's beliefs should be inextricably intertwined with her own.

12.9 Conclusion.

There are several points that we have tried to make in this paper. The first is that the simple rule (V) does not always hold; this is the negative point first made by Castañeda in 1966 but not hitherto incorporated in computational analyses of knowledge and belief. The second is a positive contribution: a single rule, implementable in logic-based representation languages such as SNePS, that can replace (V)—namely (KTB.dr) (Figure 12.5; cf. (F5.G)). Third, we demonstrated the importance of representing proper names explicitly. Fourth, we provided support for the necessity of considering sentences in the context both of extended text and of the prior beliefs that a cognitive agent understanding those sentences has in order to fully capture certain features of their semantics.[3]

12.10 Appendix.

This appendix presents sample runs (in `computer font`) of computational implementations of the two different representations for quasi-indexicals and knowledge reports. The first uses the ego case-frame, in which an incorrect inference is drawn; the second uses the 'I'-pointer, and the correct inference is drawn. Explanatory comments (in normal font) have been added. After SNePS is invoked, user input follows the '`:`' prompt, and subsequent lines show Cassie's output (some irrelevant information has been deleted or edited, for ease of readability). Cassie interprets each English sentence in terms of a SNePS network, which she then expresses in English. (For details, see Shapiro 1982; Shapiro & Rapaport 1987, 1995; Rapaport 1988b, 1991a.)

12.10.1 Demo 1.

```
: For every agnt and prop if agnt knows prop then prop.
I understand that for every agnt and prop
if agnt knows prop then prop.
Time (sec.): 0.37
```

This is rule (V), represented by node `M2!` (see Figure 12.26); informally: ($\forall$`V3, V4`)[if Know(`V3, V4`), then `V4`]. ("I understand that" is a canned phrase prepended to the SNePS/GATN-generated English expression of the network built as a result of interpreting the input sentence.)

```
: You are poor.
I understand that I am poor.
Time (sec.): 0.24
```

[3]This material is based upon work supported by the National Science Foundation under Grant Numbers IST-8504713 and IRI-8610517. An earlier version appeared as Rapaport, Shapiro, & Wiebe 1986. We are grateful to our colleagues in the SNePS Research Group (especially Sandra M. Peters) and the Discourse and Narrative Research Group of the SUNY Buffalo Center for Cognitive Science for discussions on these issues, to John Weber and Davin Milun for help with the figures, and to Afzal Ballim and two anonymous referees for very helpful comments on an earlier version of this paper. This chapter originally appeared in *Cognitive Science* 21 (1997) 63–107, ©1997, Cognitive Science Society, and is reprinted here with permission.

Cassie interprets the user's use of 'you' as referring to her, represented by an `ego` arc pointing to the node (as it happens, `B1`) that represents herself; the GATN has her express this node as 'I'.

```
: John knows that he is rich.
```

This triggers rule `M2 !`, and Cassie reasons as follows: since agent `B2` (i.e., John) knows that he* is `B3` (expressed as "`* is`"), she can infer—incorrectly (see §12.6)—that Ego(`B3`) (expressed as "`B3 is`"), and since `B2` (i.e., John) knows that `B3` is rich, Cassie infers that `B3` is rich:

```
Since for every agnt and prop if agnt knows prop then prop.
and John knows that * is.
I infer B3 is.

Since for every agnt and prop if agnt knows prop then prop.
and John knows that * is rich.
I infer I am rich.
```

The result, which has Cassie incorrectly taking herself to be `B3`, is expressed in English as follows:

```
 I understand that I am rich
and
 John knows that * is rich.
 Time (sec.): 1.08
```

Figure 12.26 shows the contents of Cassie's mind after this interaction has taken place.

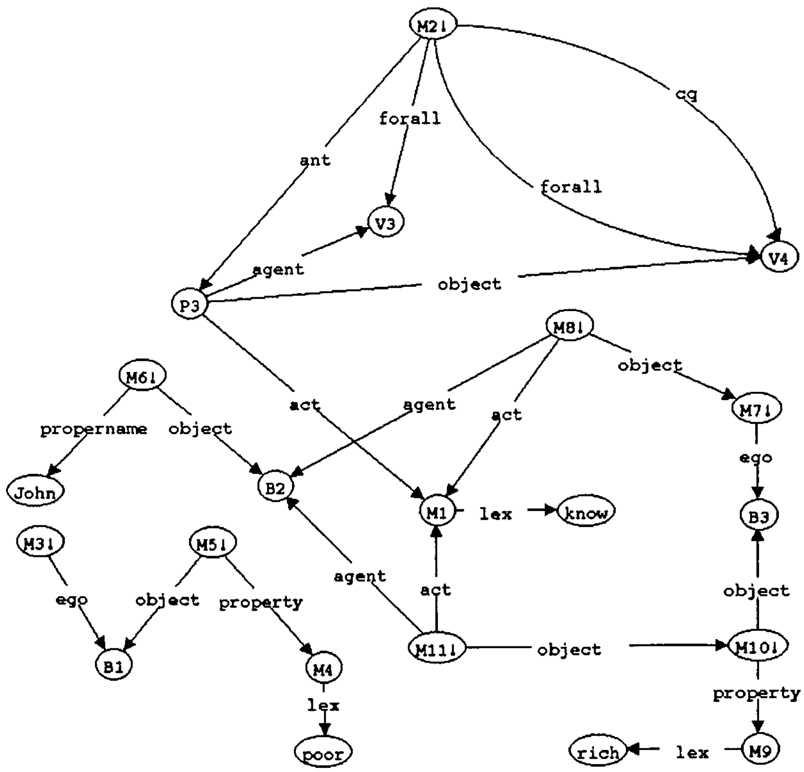

Figure 12.26: Cassie's mind after the interaction of Demo 1. The node labels in this figure and in Figure 12.27 were generated automatically by SNePS during the natural-language parsing of the input sentences; hence, they are not necessarily consistent with those in previous figures or with each other.

12.10.2 Demo 2.

The English interaction that follows is the same as before, but the SNePS representation is different; see Figure 12.27.

```
 : For every agnt and prop if agnt knows prop then prop.
 I understand that for every agnt and prop
 if agnt knows prop then prop.
 Time (sec.): 0.99

 : You are poor.
 I understand that I am poor.
 Time (sec.): 0.2

 : John knows that he is rich.
```

M8! represents the proposition that B2 (i.e., John) knows that he* (i.e., B2) is rich.
M7! represents the proposition that B2 (i.e., John) is rich.

```
Since for every agnt and prop if agnt knows prop then prop.
and John knows that * is rich.
I infer John is rich.
 I understand that John is rich
and
 John knows that * is rich.
 Time (sec.): 0.79
```

Figure 12.27 shows the contents of Cassie's mind after this interaction has taken place.

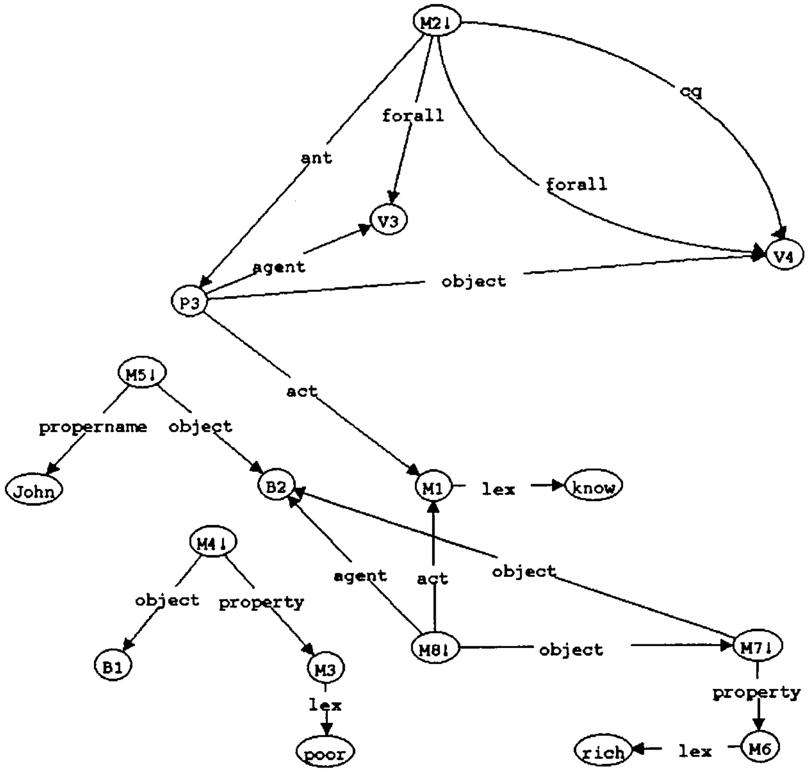

Figure 12.27: Cassie's mind after the interaction of Demo 2. Since the 'I'-pointer is implemented as a SNePS mechanism for keeping track of node B1, it is not part of the network; hence, it does not appear here.

References

Almeida, Michael J. (1987), "Reasoning about the Temporal Structure of Narratives," *Technical Report 87-10* (Buffalo: SUNY Buffalo Department of Computer Science).

Almeida, Michael J. (1995), "Time in Narratives," in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 159–189.

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

Ballim, Afzal, & Wilks, Yorick (1991), *Artificial Believers: The Ascription of Belief* (Hillsdale, NJ: Lawrence Erlbaum Associates).

Banfield, Ann (1982), *Unspeakable Sentences: Narration and Representation in the Language of Fiction* (Boston: Routledge & Kegan Paul).

Barwise, Jon, & Perry, John (1983), *Situations and Attitudes* (Cambridge, MA: MIT Press).

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 Department of Computer Science).

Castañeda, Hector-Neri (1966), " 'He': A Study in the Logic of Self-Consciousness," *Ratio* 8: 130–157.

Castañeda, Hector-Neri (1967), "Indicators and Quasi-Indicators," *American Philosophical Quarterly* 4: 85–100.

Castañeda, Hector-Neri (1970), "On the Philosophical Foundations of the Theory of Communication: Reference," in Peter A. French, Theodore E. Uehling, Jr., Howard K. Wettstein (eds.), *Studies in the Philosophy of Language*, Midwest Studies in Philosophy, Vol. 2 (Morris, MN: University of Minnesota at Morris): 165–186.

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.

Castañeda, Hector-Neri (1989), *Thinking, Language, and Experience* (Minneapolis: University of Minnesota Press).

Chalupsky, Hans (1993), "Using Hypothetical Reasoning as a Method for Belief Ascription," *Journal of Experimental and Theoretical Artificial Intelligence* 5: 119–133.

Fludernik, Monika (1993), *The Fictions of Language and the Languages of Fiction: The Linguistics Representation of Speech and Consciousness* (London: Routledge).

Fodor, Janet Dean (1970/1979), *The Linguistic Description of Opaque Contexts* (New York: Garland Publishing, 1979).

Franchere, Ruth (1964), *Hannah Herself* (New York: Thomas Y. Crowell).

Galbraith, Mary (1995), "Deictic Shift Theory and the Poetics of Involvement in Narrative," in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 19–59.

Gettier, Edmund L. (1963), "Is Justified True Belief Knowledge?", *Analysis* 23: 121–23; reprinted in A. P. Griffiths (ed.), *Knowledge and Belief* (Oxford: Oxford University Press, 1967).

Grove, Adam J. (1995), "Naming and Identity in Epistemic Logic; Part II: A First-Order Logic for Naming," *Artificial Intelligence* 74: 311–350.

Haas, Andrew R. (1993), "Indexical Expressions in the Scope of Attitude Verbs," *Computational Linguistics* 19: 637–649.

Halpern, Joseph Y. (ed.) (1986), *Theoretical Aspects of Reasoning About Knowledge: Proceedings of the 1986 Conference (Monterey, CA)* (Los Altos, CA: Morgan Kaufmann).

Hamburger, Käte (1973), *The Logic of Literature* M. J. Rose (trans.) (Bloomington: Indiana University Press).

Hendrix, Gary (1979), "Encoding Knowledge in Partitioned Networks," in Nicholas V. Findler (ed.), *Associative Networks: The Representation and Use of Knowledge by Computers* (New York: Academic Press): 51–92.

Hintikka, Jaakko (1962), *Knowledge and Belief: An Introduction to the Logic of the Two Notions* (Ithaca: Cornell University Press).

Hobbs, Jerry R. (1985), "Ontological Promiscuity," *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics (University of Chicago)* (Morristown, NJ: Association for Computational Linguistics): 61–69.

Israel, David (1989), Review of William J. Rapaport, "Logical Foundations for Belief Representation," *Journal of Symbolic Logic* 54: 617–618.

Kamp, Hans, & Reyle, Uwe (1993), *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory* (Dordrecht, Holland: Kluwer Academic Publishers).

Kuroda, S.-Y. (1976), "Reflections on the Foundations of Narrative Theory—From a Linguistic Point of View," in Teun A. van Dijk (ed.), *Pragmatics of Language and Literature* (Amsterdam: North-Holland): 107–140.

Lehnert, W. G.; Dyer, M. G.; Johnson, P. N.; Yang, C. J.; & Harley, S. (1983), "BORIS—An Experiment in In-Depth Understanding of Narratives," *Artificial Intelligence* 20: 15–62.

Lespérance, Yves (1989), "A Formal Account of Self-Knowledge and Action," in N. S. Sridharan (ed.), *Proceedings of the 11th International Joint Conference on Artificial Intelligence (IJCAI-89, Detroit)* (San Mateo, CA: Morgan Kaufmann), Vol. 2, pp. 868–874.

Lespérance, Yves, & Levesque, Hector J. (1990), "Indexical Knowledge in Robot Plans," *Proceedings of the 8th National Conference on Artificial Intelligence (AAAI-90)* (Menlo Park, CA: AAAI Press/MIT Press): 1030–1037.

Lespérance, Yves, & Levesque, Hector J. (1995), "Indexical Knowledge and Robot Action—A Logical Account," *Artificial Intelligence* 73: 69–115.

Li, Naicong (1991), "Perspective-Taking in Mandarin Discourse," Ph.D. dissertation, SUNY Buffalo Department of Linguistics.

Li, Naicong, & Zubin, David A. (1995), "Discourse Continuity and Perspective-Taking," in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 287–307.

Maida, Anthony S., & Shapiro, Stuart C. (1982), "Intensional Concepts in Propositional Semantic Networks," *Cognitive Science* 6: 291–330; reprinted in Ronald J. Brachman & Hector J. Levesque (eds.), *Readings in Knowledge Representation* (Los Altos, CA: Morgan Kaufmann, 1985): 169–189.

Martins, João, & Shapiro, Stuart C. (1988), "A Model for Belief Revision," *Artificial Intelligence* 35: 25–79.

Meinong, Alexius (1904), "Über Gegenstandstheorie," in Rudolf Haller (ed.), *Alexius Meinong Gesamtausgabe*, Vol. II (Graz, Austria: Akademische Druck- u. Verlagsanstalt, 1971): 481–535; English translation ("The Theory of Objects") by Isaac Levi et al., in Roderick M. Chisholm (ed.), *Realism and the Background of Phenomenology* (New York: Free Press, 1960): 76–117.

Plato, *Theaetetus*, in E. Hamilton & H. Cairns (eds.), *The Collected Dialogues of Plato* (Princeton: Princeton University Press, 1961): 845–919.

Quine, W. V. O. (1956/1976), "Quantifiers and Propositional Attitudes," in *The Ways of Paradox and Other Essays, Revised and Enlarged Edition* (Cambridge, MA: Harvard University Press, 1976): 185–196.

Rapaport, William J. (1978), "Meinongian Theories and a Russellian Paradox," *Noûs* 12: 153–180; errata, *Noûs* 13 (1979) 125.

Rapaport, William J. (1981), "How to Make the World Fit Our Language: An Essay in Meinongian Semantics," *Grazer Philosophische Studien* 14: 1–21.

Rapaport, William J. (1985), "Meinongian Semantics for Propositional Semantic Networks," *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics (University of Chicago)* (Morristown, NJ: Association for Computational Linguistics): 43–48.

Rapaport, William J. (1985/1986), "Non-Existent Objects and Epistemological Ontology," *Grazer Philosophische Studien* 25/26: 61–95; reprinted in Rudolf Haller (ed.), *Non-Existence and Predication* (Amsterdam: Rodopi, 1986).

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

Rapaport, William J. (1988a), Review of Halpern (ed.), *Theoretical Aspects of Reasoning About Knowledge*, in *Journal of Symbolic Logic* 53: 660–670.

Rapaport, William J. (1988b), "Syntactic Semantics: Foundations of Computational Natural-Language Understanding," in James H. Fetzer (ed.) *Aspects of Artificial Intelligence* (Dordrecht, Holland: Kluwer Academic Publishers): 81–131; reprinted in Eric Dietrich (ed.) (1994), *Thinking Computers and Virtual Persons: Essays on the Intentionality of Machines* (San Diego: Academic Press): 225–273.

Rapaport, William J. (1991a), "Predication, Fiction, and Artificial Intelligence," *Topoi* 10: 79–111.

Rapaport, William J. (1991b), "Meinong, Alexius; I: Meinongian Semantics," in Hans Burkhardt & Barry Smith (eds.), *Handbook of Metaphysics and Ontology* (Munich: Philosophia Verlag): 516–519.

Rapaport, William J. (1992), "Belief Representation Systems," in Stuart C. Shapiro (ed.), *Encyclopedia of Artificial Intelligence, 2nd edition* (New York: John Wiley): 98–110.

Rapaport, William J. (1995), "Understanding Understanding: Syntactic Semantics and Computational Cognition," in James E. Tomberlin (ed.), *Philosophical Perspectives*, Vol. 9: *AI, Connectionism, and Philosophical Psychology* (Atascadero, CA: Ridgeview): 49–88.

Rapaport, William J.; Segal, Erwin M.; Shapiro, Stuart C.; Zubin, David A.; Bruder, Gail A.; Duchan, Judith F.; Almeida, Michael J.; Daniels, Joyce H.; Galbraith, Mary M.; Wiebe, Janyce M.; & Yuhan, Albert Hanyong (1989a), "Deictic Centers and the Cognitive Structure of Narrative Comprehension," *Technical Report 89-01* (Buffalo: SUNY Buffalo Department of Computer Science).

Rapaport, William J.; Segal, Erwin M.; Shapiro, Stuart C.; Zubin, David A.; Bruder, Gail A.; Duchan, Judith F.; & Mark, David M. (1989b), "Cognitive and Computer Systems for Understanding Narrative Text," *Technical Report 89-07* (Buffalo: SUNY Buffalo Department of Computer Science).

Rapaport, William J., & Shapiro, Stuart C. (1984), "Quasi-Indexical Reference in Propositional Semantic Networks," *Proceedings of the 10th International Conference on Computational Linguistics (COLING-84, Stanford University)* (Morristown, NJ: Association for Computational Linguistics): 65–70.

Rapaport, William J.; Shapiro, Stuart C.; & Wiebe, Janyce M. (1986), "Quasi-Indicators, Knowledge Reports, and Discourse," *Technical Report 86-15* (Buffalo: SUNY Buffalo Department of Computer Science).

Reboul, Anne (1992), "How Much Am *I* I and How Much is *She* I?", *Lingua* 87: 169–202.

Roberts, Lawrence D., & Rapaport, William J. (1988), "Quantifier Order, Reflexive Pronouns, and Quasi-Indexicals," *Technical Report 88-16* (Buffalo: SUNY Buffalo Department of Computer Science).

Russell, Bertrand (1905), "On Denoting," reprinted in *Logic and Knowledge: Essays 1901–1950*, R. C. Marsh (ed.) (New York: Capricorn, 1956): 39–56.

Seager, William (1990), "The Logic of Lost Lingens," *Journal of Philosophical Logic* 19: 407–428.

Segal, Erwin M. (1995), "Narrative Comprehension and the Role of Deictic Shift Theory," in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 3–17.

Sells, Peter (1987), "Aspects of Logophoricity," *Linguistic Inquiry* 18: 445–479.

Shapiro, Stuart C. (1975), "Generation as Parsing from a Network into a Linear String," *American Journal of Computational Linguistics*, Microfiche 33: 45–62.

Shapiro, Stuart C. (1979a), "Generalized Augmented Transition Network Grammars for Generation from Semantic Networks," *Proceedings of the 17th Annual Meeting of the Association for Computational Linguistics (University of California at San Diego)*: 25–29.

Shapiro, Stuart C. (1979b), "The SNePS Semantic Network Processing System," in Nicholas V. Findler (ed.), *Associative Networks: The Representation and Use of Knowledge by Computers* (New York: Academic Press): 179–203.

Shapiro, Stuart C. (1982), "Generalized Augmented Transition Network Grammars for Generation from Semantic Networks," *American Journal of Computational Linguistics* 8: 12–25.

Shapiro, Stuart C. (1990), "Cables, Paths, and 'Subconscious' Reasoning in Propositional Semantic Networks," in John Sowa (ed.), *Formal Aspects of Semantic Networks* (Los Altos, CA: Morgan Kaufmann): 137–156.

Shapiro, Stuart C. (1993), "Belief Spaces as Sets of Propositions," *Journal of Experimental and Theoretical Artificial Intelligence* 5: 225–235.

Shapiro, Stuart C., & Martins, João P. (1990), "Recent Advances and Developments: The SNePS 2.1 Report," in Deepak Kumar (ed.), *Current Trends in SNePS*, Lecture Notes in Artificial Intelligence No. 437 (Berlin: Springer-Verlag): 1–13.

Shapiro, Stuart C., & Rapaport, William J. (1987), "SNePS Considered as a Fully Intensional Propositional Semantic Network," in Nick Cercone & Gordon McCalla (eds.), *The Knowledge Frontier: Essays in the Representation of Knowledge* (New York: Springer-Verlag): 262–315.

Shapiro, Stuart C., & Rapaport, William J. (1991), "Models and Minds: Knowledge Representation for Natural-Language Competence," in Robert Cummins & John Pollock (eds.), *Philosophy and AI: Essays at the Interface* (Cambridge, MA: MIT Press): 215–259.

Shapiro, Stuart C., & Rapaport, William J. (1992), "The SNePS Family," *Computers and Mathematics with Applications* 23: 243–275; re-printed in F. Lehmann (ed.), *Semantic Networks in Artificial Intelligence* (Oxford: Pergamon Press, 1992): 243–275.

Shapiro, Stuart C., & Rapaport, William J. (1995), "An Introduction to a Computational Reader of Narratives," in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 79–105.

Shapiro, Stuart C., & the SNePS Implementation Group (1994), "SNePS-2.1 User's Manual," *SNeRG Technical Note* (Buffalo: SUNY Buffalo Department of Computer Science).

Tomberlin, James E. (ed.) (1983), *Agent, Language, and the Structure of the World* (Indianapolis: Hackett).

Tomberlin, James E. (ed.) (1986), *Hector-Neri Castañeda* (Dordrecht, Holland: D. Reidel).

Von Eckardt, Barbara (1993), *What is Cognitive Science?* (Cambridge, MA: MIT Press).

Wiebe, Janyce M. (1990a), "Recognizing Subjective Sentences: A Computational Investigation of Narrative Text," *Technical Report 90-03* (Buffalo: SUNY Buffalo Department of Computer Science).

Wiebe, Janyce M. (1990b), "Identifying Subjective Characters in Narrative," *Proceedings of the 13th International Conference on Computational Linguistics (COLING-90, University of Helsinki)* (Association for Computational Linguistics), Vol. 2, pp. 401–408.

Wiebe, Janyce M. (1991), "References in Narrative Text," *Noûs* 25: 457–486; reprinted in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates, 1995): 263–286.

Wiebe, Janyce M. (1994), "Tracking Point of View in Narrative," *Computational Linguistics* 20: 233–287.

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.

Wiebe, Janyce M., & Rapaport, William J. (1988), "A Computational Theory of Perspective and Reference in Narrative" *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics (SUNY Buffalo)* (Morristown, NJ: Association for Computational Linguistics): 131–138.

Wilks, Yorick; Ballim, Afzal; & Dietrich, Eric (1989), "Pronouns in Mind: Quasi-Indexicals and the 'Language of Thought'," *Computers and Artificial Intelligence* 8: 493–503.

Wyatt, Richard (1989), "The Representation of Opaque Contexts," *Technical Report 89-13* (Buffalo: SUNY Buffalo Department of Computer Science).

Wyatt, Richard (1990), "Kinds of Opacity and Their Representations," in Deepak Kumar (ed.), *Current Trends in SNePS*, Lecture Notes in Artificial Intelligence No. 437 (Berlin: Springer-Verlag): 123–144.

Wyatt, Richard (1993), "Reference and Intensions," *Journal of Experimental and Theoretical Artificial Intelligence* 5: 263–271.