THOUGHT, LANGUAGE, AND ONTOLOGY:
Essays in Memory of Hector-Neri Castañeda

Edited by

Francesco Orilia
Dipartimento di filosofia e scienze umane
Università di Macerata, 62100 Macerata, Italy
orilia@unimc.it

and

William J. Rapaport
Department of Computer Science, Department of Philosophy,
and Center for Cognitive Science
State University of New York at Buffalo, Buffalo, NY 14260-2000, USA
rapaport@cs.buffalo.edu
http://www.cs.buffalo.edu/pub/WWW/faculty/rapaport/

Prolegomena to a Study of Hector-Neri Castañeda's Influence on Artificial Intelligence: A Survey and Personal Reflections

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

C.1 Introduction.

In 1982, I made the transition from being a professional philosopher to being a professional computer scientist and "intelligence artificer" (to use Daniel Dennett's happy term)—"professional" in the sense that that is now how I *earn* my living, though not in the sense that that is how I live my professional life—for my philosophical and artificial-intelligence (AI) research have dovetailed so well that I am hard pressed to say where one leaves off and the other begins. I remember Castañeda telling me at the time that he, too, felt that philosophy and AI were intimately related—that the importance of AI lay in the fact that it filled in—indeed, *had* to fill in—all the gaps left in abstract philosophical theories; it was in AI that all the 'i's were dotted and 't's crossed, since AI programs had to be executable and could not leave anything to be specified at a later time. Thus, for Castañeda, AI would keep philosophers honest, while philosophy could provide AI with ideas and theories to be implemented.

Hector-Neri Castañeda's own philosophical ideas and theories have influenced AI researchers, and have the potential for even more such influence. The actual influence has taken both direct and indirect forms: direct, in that a handful of AI researchers have referred explicitly to his work and, more importantly, have incorporated it into their own; indirect, in that, first, several AI researchers have been influenced by philosophers and other AI workers who were themselves directly influenced by Castañeda, and, second, several of Castañeda's students have joined me in making AI the focus of our own work.

C.2 Actual Influence.

C.2.1 Direct Influences.

Let me begin with a brief survey of the AI work that has been directly inspired by two of Castañeda's major philosophical contributions: his theory of intentions, practitions, and the ought-to-do, and his theory of quasi-indicators.

C.2.1.1 Intentions, practitions, and the ought-to-do.

It is interesting to note that almost all the major subfields of AI mirror subfields of philosophy: The AI analogue of philosophy of language is computational linguistics; what philosophers call "practical reasoning" is called "planning and acting" in AI;[1] ontology (indeed, much of metaphysics and epistemology) corresponds to knowledge representation in AI; and automated reasoning is one of the AI analogues of logic.

C.2.1.1.1 L. Thorne McCarty. Deontic logic finds its AI home in several different, though related, areas: knowledge representation and reasoning as well as planning and acting. L. Thorne McCarty, of the Rutgers computer science department, was one of the first researchers to introduce deontic concepts such as permission and obligation into AI, where they are needed for representing and reasoning about legal concepts, planning and acting, and even computer security (cf. the notions of "read permission" vs. "write permission"). In "Permissions and Obligations" (McCarty 1983), he sets himself the goal of developing a formal semantics for permission and obligation that avoids various deontic paradoxes in a way congenial to AI, i.e., in an implementable way. He writes:

> Instead of representing the deontic concepts as *operators* applied to *propositions*, … we will represent them as *dyadic forms* which take *condition descriptions* and *action descriptions* as their arguments. … Instead of granting permissions and imposing obligations on the state of the world itself, we will grant permissions and impose obligations on the *actions*

[1] My former colleague Randall R. Dipert was the first to point this out to me.

> which *change* the state of the world. This is an approach long advocated by Castañeda …. (McCarty 1983: 288.)

Here, McCarty cites Castañeda's 1975 book, *Thinking and Doing*, and his 1981 article, "The Paradoxes of Deontic Logic". In a 1986 sequel to his paper, McCarty elaborates on this:

> The deontic expressions should distinguish the *condition* part of a rule from the *action* part of a rule, instead of treating these as similar expressions linked by logical implication. The use of a distinct syntactic condition here is not a new idea; it is the central feature of all systems of dyadic deontic logic [here, he cites David Lewis]. However, if we combine this idea of a distinct syntactic condition with the previous idea of a distinct action language, we have a system which includes both 'propositions' and 'practitions' in the sense of Hector Castañeda [again, he cites *Thinking and Doing* and "The Paradoxes of Deontic Logic"]. Castañeda has long argued that the failure to make this fundamental distinction between 'propositions' and 'practitions' is the source of most of the deontic paradoxes." (McCarty 1986: 5–6.)

It's natural that McCarty should cite Castañeda, since he's working on deontic logic, Castañeda is an important figure in deontic logic, and it's a research strategy of AI to look at the philosophical (as well as psychological, linguistic, etc.) literature for relevant work. It is important, however, to note that McCarty *adds* something to the mix, and doesn't just appropriate Castañeda's theory. In particular, McCarty says that

> to carry out this approach in full [namely, the one "long advocated by Castañeda"] it seems necessary to establish a connection between the abstract description of an action and the concrete changes that occur in the world when the action takes place. This has been a major concern of artificial intelligence research …. (McCarty 1983: 288.)

Thus, input from AI serves to flesh out the philosophical theory, giving rise to mutual interaction between AI and philosophy, in much the way Castañeda advocated.

Upon more careful examination, however, we find that it is not Castañeda's proposition–practition distinction *per se* that McCarty builds on. For Castañeda, a practition has the form "agent *A* to do action *X*". Rather, McCarty uses what might be generalized as a proposition/other-than-proposition distinction: For what plays the role of Castañeda's practitions in McCarty's theory are what McCarty calls (somewhat confusingly, given the present context) "actions", where a primitive action is a state change—a change in the world from a situation in which state $S1$ is true to a situation in which state $S2$ is true (McCarty 1983: 290), where $S1$ and $S2$ range over formulas of forms such as "actor *a* stands in relation *r* to object *o*".

C.2.1.1.2 Philip R. Cohen and Hector J. Levesque. An important interdisciplinary anthology titled *Intentions in Communication*, edited in 1990 by two AI researchers (Philip R. Cohen, of the AI Center at SRI International, and Martha E. Pollack, now of the University of Pittsburgh computer science department) and a linguist (Jerry Morgan), cites Castañeda in several places. The editor's introduction (Cohen, Morgan, & Pollack 1990b) notes that "[a]nalyses of the nature of intention are plentiful in the philosophy of action" (p. 2), citing Castañeda's *Thinking and Doing* in the same breath as such philosophers as G. E. M. Anscombe, Robert Audi, Michael Bratman, Donald Davidson, and Alvin Goldman, inter alia. And Cohen and Hector J. Levesque (of the University of Toronto computer science department) write in their contribution, "Persistence, Intention, and Commitment" (1990), that

> Intention has often been analyzed differently from other mental states such as belief and knowledge. First, whereas the content of beliefs and knowledge is usually considered to be in the form of propositions, the content of an intention is typically regarded as an action. For example, Castañeda [here they cite *Thinking and Doing*] treats the content of an intention as a "practition," akin (in computer science terms) to an action description. It is claimed that by doing so, and by strictly separating the logic of propositions from the logic of practitions, one avoids undesirable properties in the logic of intention, such as the fact that if one intends to do an action a, one must also intend to do a or b. However, it has also been argued that needed connections between propositions and practitions may not be derivable (Bratman 1983). (Cohen & Levesque 1990: 35.)

Bratman, too, cites Castañeda in his contribution to this volume, making an interesting observation on the relevance of Castañeda's contribution to AI:

> A contrasting literature—best represented by Hector-Neri Castañeda's book *Thinking and Doing* …—has emphasized the role of prior intentions as inputs into practical reasoning. The paradigm here, broadly speaking, is reasoning from a prior intention and relevant beliefs to derivative intentions concerning means, preliminary steps, or more specific courses of action. My close advisors on matters of artificial intelligence—David Israel and Martha Pollack—have taught me that this is also the paradigm that dominates there as well, in the guise of the "planning problem." (Bratman 1990: 17.)

As I will note later, Bratman, although a philosopher and not an AI researcher, has been very important in getting Castañeda's message to the AI community.

C.2.1.1.3 Martha E. Pollack. Pollack also alludes to Castañeda's notion of practitions and intentions (which are first-person practitions, of the form "I to do *X*" (Castañeda 1975: 172)) in her 1991 essay, "Overloading Intentions for Efficient Practical Reasoning". Pollack is concerned with planning in dynamic environments—ones that can change before the planning is finished. She suggests a way to achieve two unrelated goals when an action intended to achieve one can also help achieve the other, and she shows how her method can be more efficient in terms of reasoning and action than the "decision-theoretic model in which a complete set of alternatives is first generated and then weighed against one another," even though the decision-theoretic model might produce a more "optimal" plan (Pollack 1991: 521). She sets up the following example:

> My expectation that Steve will pass by my office may be central in the plan I adopt to get money for lunch: I may decide to borrow it from him when he passes by. (Pollack 1991: 530.)
>
> … when an agent exploits a secondary expectation, she forms, without complete deliberation, a new intention to perform the action that she had originally merely expected to perform as a side-effect of some other action. When I decide to make a withdrawal using the ATM, expecting that I will pass the ATM, I form the intention to pass the ATM. But when an agent forms a plan that relies on an independent expectation, she does not, in general, form an intention to bring about that expectation.[7] It seems unlikely, in the current example, that I would form an intention to ensure that Steve passes by my office shortly before noon. (Pollack 1991: 531.)

Her note 7 refers to an argument "that it may often be impossible for an agent to form an intention whose object … is the independent expectation" (Pollack 1991: 535). In particular, she goes on, "… I cannot form an intention whose object is the action of Steve's passing by the office shortly before noon; as Castañeda has argued, intentions are first-person attitudes"; here, she cites *Thinking and Doing* (Pollack 1991: 535). Pollack's work is part of a research program in philosophy and AI, with insights and results from each providing data for the other. As she notes at the end of her essay, "philosophical theories can and do matter to the AI researcher" (Pollack 1991: 534).

In response to a query, she informs me that "work on joint intentions done by [Barbara J.] Grosz [of the Harvard computer science department], … in collaboration with [Candace L.] Sidner [a computational linguist with Lotus Development Corp.] … probably was influenced by Castañeda, as one of the things they worked hard to do was preserve the first-person nature of intentions in their model" (personal communication, 26 September 1995; she cites Grosz & Sidner 1990, Grosz & Kraus 1993).

Grosz confirms this (personal communication, 28 November 1995); she writes:

> We don't cite Castañeda. Much to my embarrassment I have not read his work (it's on the proverbial stack). I suspect the influence is indirect:

> Martha is exactly right that one of the things [we] worked hard to do was preserve the first-person nature of intentions. This need was certainly driven home to me in discussions with Bratman, and perhaps other philosophers at Stanford/CSLI. I think by the time I heard talks on Castañeda this was already a goal. And it had other sources (I never liked SIHI[2] formulations of speech acts).

C.2.1.2 Ought-to-do vs. ought-to-be.

A query to the `comp.ai` electronic bulletin-board newsgroup on Castañeda's contributions to AI brought a response from Christen Krogh, a computer-scientist-*turned-philosopher* at the University of Oslo. In "Getting Personal" (Krogh 1996), he discusses Castañeda's distinction between the ought-to-do and the ought-to-be (citing Castañeda's 1970 essay "On the Semantics of the Ought-to-Do"), and applies it to what he calls "personal ought-to-do … expressions" (Krogh 1996, §9), i.e., expressions of the form "it is obligatory for [person] *i* that *i* sees to it that *A*" (Krogh 1996, §§3, 9). Curiously, Krogh does not cite Castañeda's *Thinking and Doing*, in which the notions of intention and practition would seem to do the work that Krogh wants.

In any event, I was about to despair that these few examples of the influence of Castañeda's theories of deontic logic were the only ones, when I obtained—at Krogh's suggestion—a copy of an anthology of papers for DEON'91, *Deontic Logic in Computer Science*, edited by John-Jules Ch. Meyer (a computer scientist at Utrecht University) and Roel J. Wieringa (a computer scientist at Vrije Universiteit in Amsterdam). There, between the copyright page and the table of contents, appears in large, boldface type "**In Memoriam Hector-Neri Castañeda**"! The editors explain in their Preface:

> Finally, we mention that originally also H.-N. Castañeda (Indiana University) had accepted an invitation to present a paper at DEON'91, which paper undoubtedly would have been included in this volume, but regrettably he died some months before the conference. His vast and original work in deontic logic in particular and philosophy in general will always remain a source of inspiration. (Meyer & Wieringa 1993a: xii.)

And these are, recall, the words of two *computer scientists.* The volume contains an article by them referring to Castañeda's 1981 "The Paradoxes of Deontic Logic" as well as a paper analyzing Castañeda's contributions by the philosopher Risto Hilpinen (Meyer & Wieringa 1993b, Hilpinen 1993).

[2] I (Rapaport) believe this stands for "Speaker Intention/Hearer Intention".

C.2.1.3 Quasi-indicators and intensionality.

I now turn to the second major philosophical theory of Castañeda's that has been directly adopted into AI: the theory of quasi-indicators. A quasi-indicator is 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. Castañeda's discovery of quasi-indicators in English is probably his most widely accepted contribution to philosophy. The additional facts—as the French linguist and philosopher Anne Reboul[3] and I have independently noted (Reboul 1992; Rapaport, Shapiro, & Wiebe 1997)—that some languages have distinct expressions for quasi-indicators (called "logophoric pronouns"; cf. Sells 1987) and that so-called "free indirect discourse" in literature uses quasi-indicators to represent a character's thoughts (cf. Wiebe & Rapaport 1988, and Wiebe 1990a, 1990b, 1991, 1994 for discussion and further references) make it natural to expect that computational linguists and knowledge-representation researchers would have to incorporate quasi-indicators in their theories.

C.2.1.3.1 Andrew R. Haas. One such computational linguist is Andrew R. Haas, of SUNY Albany's computer science department. In a 1993 paper in *Computational Linguistics*, called "Intensional Expressions in the Scope of Attitude Verbs,"—note that this title would be equally at home in a philosophy or a pure linguistics journal—he considers "a sentential theory of attitudes," i.e., one that "holds that propositions (the things that agents believe and know) are sentences of a representation language," arguing that the propositions expressed by utterances are *not* sentences but singular Kaplanesque propositions (Kaplan 1989). He then "shows how such a theory can describe the semantics of attitude verbs and account for the opacity of indexicals in the scope of these verbs" (Haas 1993: 637).

In particular, he applies his theory to quasi-indicators. Haas asks us to

> Suppose John says "I am smart." Hearing his words, one would naturally describe John's belief by saying "He thinks he is smart." In this sentence the pronoun "he" appears in the scope of an attitude verb, and it represents the subject's use of a first person pronoun. Castañeda [here, he cites the 1968 essay "On the Logic of Attributions of Self-Knowledge to Others"] coined the term **quasi-indicator** for an occurrence of a pronoun that is in the scope of an attitude verb and that represents the subject's use of some indexical expression. (Haas 1993: 646.)

His analysis requires a "selfname" for each agent. This is "a standard constant" used by the speaker "to refer to himself or herself" (Haas 1993: 646). Thus, presumably, an agent John's own representation of his utterance "I am smart" would be something

[3] Who currently works on AI in a computer-science lab.

like:

smart(i)

or, perhaps

believe(i, smart(i))

where 'i' is John's selfname. A selfname, of course, is not a quasi-indicator, since a quasi-indicator is an expression used by *another* speaker to represent someone else's selfname. To represent sentences containing quasi-indicators, 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. (Haas 1993: 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).

Haas closes with remarks about how these are important constraints in ordinary language, hence that computational analyses must be able to handle them.[4]

C.2.1.3.2 Yves Lespérance and Hector J. Levesque. Two other AI researchers who feel the same way are Levesque and Yves Lespérance, also of the University of Toronto computer science department (Lespérance 1989; Lespérance & Levesque 1990, 1995). In their theory of indexical knowledge and robot action, published in *Artificial Intelligence*—the most prestigious journal for AI research—Lespérance and Levesque introduce two terms, "**self**, which denotes the current agent, and **now**, which denotes the current time" (1995: 80):

> In English, there are true indexicals (I, you, now, here, etc.), which refer to aspects of the utterance context no matter where they appear, and there

[4] In a personal communication (14 February 1996), Haas observes that "[i]n computational linguistics we are hearing a great deal about "empirical" work: extracting linguistic information automatically from large samples of text, instead of encoding it by hand. Even people who believe in hand-encoded knowledge probably welcome the new emphasis on data-collection and performance measurement. If AI is going to become more empirical, it will sharpen the contrast with philosophy, which is usually not empirical in any systematic way. The word "systematic" is crucial there—philosophers can be very acute observers of human behavior, and many important observations have first appeared in papers on philosophy of language. Yet if you collect data systematically, you are a linguist not a philosopher." The present essay, he goes on, "comes close to these issues in the first paragraph, where … Castañeda [is quoted] about dotting i's and crossing t's. … Why isn't philosophy of language more empirical? Are the philosophers making a mistake? Or is data-collection not relevant to their goals? If not, there is a large difference between their goals and ours …. Montague urged his colleagues to practice "formal philosophy". Is "empirical philosophy" coming? Or is it a contradiction in terms?"

> are quasi-indexicals/quasi-indicators [here, they cite Castañeda's "On the Logic of Attributions of Self-Knowledge to Others"] (I myself, you yourself, he himself, etc.), which are used to report that an agent has an indexical mental state. The behavior of our *primitive indexicals* **self** and **now** displays characteristics of both categories. 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 (1989b: 135–136), there *are* temporal quasi-indicators in English. Furthermore, Lespérance and Levesque's "primitive indexicals" 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-indicator 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-indicator. Languages with logophoric pronouns, as I noted earlier, have distinct morphemes for use in quasi-indicator contexts. (For more detailed discussion of Haas 1993 and of Lespérance & Levesque 1995, see Rapaport, Shapiro, & Wiebe 1997.)

C.2.1.3.3 The Buffalo connection. My colleague in the SUNY Buffalo computer science department, Stuart C. Shapiro, and I, along with my former student—one of Castañeda's AI *grand*students—Janyce M. Wiebe (now with the computer science department and Computing Research Laboratory at New Mexico State University) have also been developing a computational theory of quasi-indexical belief and knowledge reports.

Indeed, this was my entry ticket to AI: When I was retraining myself in computer science and seeking a research project for a master's thesis, I read an article by Shapiro and Anthony S. Maida, "Intensional Concepts in Propositional Semantic Networks," which appeared in *Cognitive Science* in 1982. There, they consider a sentence such as "John believes that he is rich," and offer a representation in the SNePS semantic-network knowledge-representation and reasoning system (Shapiro 1979; Shapiro & Rapaport 1987, 1992, 1995) that apparently ignores the quasi-indexical status of that occurrence of 'he', namely, as (roughly, for I do not want to burden you with semantic-network diagrams): Believes(John, Rich(John)), where both of these occurrences of 'John' are, in the SNePS network, really one occurrence of a single term. What is not

specified in their paper is whether this single term denotes the individual named 'John' or his own representation of himself. As it happens, the first occurrence in my linear rendering here represents the individual; the second represents his own representation of himself. But Maida and Shapiro's notation incorrectly conflates these.

When I read their paper, the "Aha!" light in my mind went on: Here was my master's thesis: With my knowledge of Castañeda's theory of quasi-indicators, I could correct their representation and implement a natural-language understanding and natural-language generation system, using SNePS as its semantic theory, that would be able to understand and reason with quasi-indicators. This research was reported in our 1984 computational-linguistics conference paper, "Quasi-Indexical Reference in Propositional Semantic Networks," and my 1986 *Cognitive Science* article, "Logical Foundations for Belief Representation", where I also argued that many other AI systems needed to, but did not, pay adequate attention to Castañeda's theory. (I like to think that it was a talk on this that I gave at Toronto when Lespérance was a grad student that inspired his use of quasi-indicators!)

Wiebe and I embedded this computational theory of quasi-indicators in the broader computational theory of "Representing *De Re* and *De Dicto* Belief Reports in Discourse and Narrative" (Wiebe & Rapaport 1986; this is likely to be the strangest title ever to appear in the *Proceedings of the Institute for Electrical and Electronic Engineers*!). There we investigated the disambiguation of belief reports as they appear in discourse and narrative. In the 1986 *Cognitive Science* paper, the distinction between *de re* and *de dicto* belief reports was made solely on the basis of their representations. This analysis, however, is sufficient only when belief reports are considered in isolation. We need to consider more complicated belief structures in discourse and narrative. Further, we cannot meaningfully apply one, but not the other, of the concepts *de re* and *de dicto* to these more complicated structures. We argued that the concepts *de re* and *de dicto* do not apply to an agent's conceptual representation of her beliefs, but that they apply to the utterance of a belief report on a specific occasion. A cognitive agent interprets a belief report such as "*S* believes that *N* is *F*" (where *S* and *N* are names or descriptions and *F* is an adjective) *de dicto* if she interprets it from *N*'s perspective, and she interprets it *de re* if she interprets it from her own perspective. This, of course, is closely related to Castañeda's claims in his 1970 paper, "On the Philosophical Foundations of the Theory of Communication: Reference," which Wiebe also cites in her 1994 *Computational Linguistics* paper, "Tracking Point of View in Narrative", which is based on her 1990 dissertation, in which she solves the problem of disambiguating *de re* from *de dicto* belief reports in the special case of distinguishing subjective from objective contexts in narrative.

We have also extended our computational theory of quasi-indicators to provide a full computational analysis of *de re*, *de dicto*, and *de se* belief and knowledge reports (Rapaport, Shapiro, & Wiebe 1997). Our analysis solves a problem first observed by Castañeda in " 'He': A Study in the Logic of Self-Consciousness" (1966), namely, that the simple rule '(*A* knows that *P*) implies *P*' apparently does not hold if *P* con-

tains a quasi-indicator. We have formulated a single rule, in the context of the SNePS knowledge-representation and reasoning system, that holds for all *P*, including those containing quasi-indicators. In so doing, we explore the difference between reasoning in a public communication language and in a knowledge-representation language, the importance of representing proper names explicitly, and 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.

There is another aspect to my entry into AI that was Castañeda-inspired. My philosophy dissertation was on intensionality and the structure of existence (1976). Impressed by Castañeda's guise theory, I attempted a revision of Meinong's theory of the objects of thought that would be immune to the standard Russellian objections. Years later, in a chance encounter with Shapiro, I learned that he, too, had been working on the problems of intensionality and intentionality—that, in fact, his SNePS system was predicated on the claim that intensionality was the right way to go with building a knowledge-representation system for cognitive AI applications. (This was the main point of his paper with Maida.) Thus began our collaboration, which resulted in a Meinongian semantics for SNePS (Shapiro & Rapaport 1987) and a paper in which we argued in some detail for the necessity for, and nature of, intensional representations in AI (Shapiro & Rapaport 1991). Our central arguments—which refer to some of Castañeda's claims—are, briefly, that knowledge-representation systems intended to model (or, more strongly, to *be*) the mind of a (computational) cognitive agent have to be intensional in order to be able fully to represent and reason, first, about "fine-grained" entities, such as distinctions between the morning star and the evening star and beliefs the agent might have about the one but not the other, and, second, about "displaced" entities, such as non-existent, impossible, or fictional objects. After all, *we* are able to think about them; thus, so should computational cognitive agents be able to. Indeed, the need to be able to represent and reason about fictional entities is crucial to many AI projects designed to capture our abilities to read, understand, and produce fictional narratives, such as our project at the SUNY Buffalo Center for Cognitive Science on understanding deictic phenomena in narrative and discourse (see Rapaport 1991 and Rapaport & Shapiro 1995 for a discussion of Castañeda's theory of fictional entities in a computational context; on our deixis project, see Duchan et al. 1995; on generation, see Bringsjord 1992). I gave a more general argument for the importance for AI of Meinongian theories in general, and Castañeda's guise theory in particular, in a 1985 computational-linguistics conference paper, "Meinongian Semantics for Propositional Semantic Networks" and in "Meinongian Semantics and Artificial Intelligence," which is a contribution to *Essays on Meinong*, a perennially forthcoming book edited by Peter Simons.

C.2.2 Indirect Influences.

In addition to the direct influence of Castañeda's theories on AI just discussed, AI has also been influenced indirectly by Castañeda's theories, through the writings of philosophers such as Michael Bratman (e.g., 1987) and John Perry (e.g., 1979), who were directly influenced by Castañeda, and who, in turn, have directly influenced many AI researchers, as well as through several of Castañeda's students. Let me briefly describe some of the latter work. At least five of Castañeda's students are now working in AI:

C.2.2.1 Lewis Creary.

Lewis Creary, now at Hewlett-Packard Labs in Palo Alto, studied with Castañeda at Wayne State (where he did a master's in philosophy) and also had an opportunity to discuss computational and philosophical issues with him when Castañeda was at the Center for Advanced Study in the Behavioral Sciences near Stanford. He has written an intensional knowledge representation system for natural-language semantics and was one of the developers of HPSG (Head-driven Phrase-Structure Grammar), a major computational-linguistics grammar formalism. He has published papers in such AI forums as the *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)* and the *Proceedings of the Association for Computational Linguistics* (see Creary 1979, 1983; Creary & Pollard 1985).

C.2.2.2 Donald E. Nute.

Donald E. Nute was one of Castañeda's Ph.D. students at Indiana. Although he is chair of the philosophy department at the University of Georgia, most of his work over the last decade has been in AI. Indeed, he is also director of the AI Center at Georgia. In a reply to my query about Castañeda's influence on his AI research, he writes:

> I have often quoted Hector to my students as saying, "The best theory is always the simplest unless, of course, the world is complex." We all know how rich and complex Hector's world was!
>
> Although I studied logic with [J. Michael] Dunn, [Nino] Cocchiarella, [Robert G.] Meyer, and a couple of fellows in the math department, Hector may have actually had the most profound effect on the way I think about and do philosophical logic. I recall that his work on the logic of imperatives was criticized because the "inferences" weren't "truth-preserving". So he called them "sh-inferences" instead! Similar objections have been raised to nonmonotonic logic, an area in which I work, by Israel Scheffler. Apparently, I inherited from Hector a more catholic view of logic. I see it primarily as a tool for attacking philosophical and other problems.

> I see logical relations between propositions, imperatives, questions, policies, desires, and other entities that some logical "purists" might reject. Formal methods can be used effectively almost everywhere.
>
> So my approach to philosophical logic, as well as my work in epistemology and philosophy of language, was largely molded by Hector's influence. Ultimately, I have become involved in artificial intelligence where I continue to do epistemology, philosophy of language, and especially philosophical logic under a new guise. In particular, I work in knowledge representation, nonmonotonic logic, and expert systems. (Personal communication, 15 September 1995.)

Nute has co-authored a text on the Prolog programming language and has published articles in such AI anthologies as *Aspects of Artificial Intelligence*, *Knowledge Representation and Defeasible Reasoning*, and the *Handbook of Logic in Artificial Intelligence and Logic Programming*, as well as in such journals as *JETAI*—the *Journal of Experimental and Theoretical Artificial Intelligence*, inter alia (see Covington et al. 1988; Nute 1988ab, 1990ab, 1992, 1993, 1994, and forthcoming; Billington et al. 1990; Nute, Mann, & Brewer 1990; Potter et al. 1990; Karickhoff et al. 1991; Meyer et al. 1991; Macura et al. 1992, 1993; Geerts et al. 1994; Nute, Rauscher, et al.; Gorokhovski & Nute, forthcoming).

C.2.2.3 Robert A. Morris.

Like me, Robert A. Morris, another Indiana Ph.D. student, retrained himself as a computer scientist. He is now in the computer science department at Florida Institute of Technology and has been a visiting researcher at the Institute for the Interdisciplinary Study of Human and Machine Cognition at the University of West Florida. He informs me that:

> His research interests in AI include reasoning about time and diagnostic reasoning. He has had extensive research experience with NASA, having recently completed developing for NASA a system which applies model-based reasoning to the task of diagnosing faults to a spacecraft power distribution system. More recently, he has been involved with a research group at NASA-Ames Research Center, working on an AI system for automatic telescope observation scheduling for remote, automatic telescopes.
>
> I'm sure Hector would be proud, both of his own accomplishments, and those of his students. (Personal communication, 13 September 1995.)

Morris's AI publications have appeared in *JETAI*, IJCAI, and *Computational Intelligence* (see Morris & Marchetti 1989; Morris & Al-Khatib 1991; Morris et al. 1993, 1996; Gonzalez et al., forthcoming).

C.2.2.4 Francesco Orilia.

Francesco Orilia, who also did his Ph.D. with Castañeda at Indiana and is now in the philosophy department at the University of Macerata, Italy, writes that he

> was with the Olivetti lab in Pisa [Italy] from [19]87 to [19]94 where the focus of my activity was AI from [19]87 to [19]91. In any case AI remains among my fields of interest …. (Personal communication, 20 October 1995.)

His AI publications have appeared in *Minds and Machines: Journal for Artificial Intelligence, Philosophy, and Cognitive Science* and the Italian AI conference proceedings (see Orilia 1992, 1994ab, 1995; see also Orilia 1994c, where Cohen and Levesque's theory of intentions is criticized from the vantage point of Castañeda's practition/proposition distinction).

In particular, his 1994 *Minds and Machines* paper, "Belief Representation in a Deductivist Type-Free Doxastic Logic", is concerned with what AIers call 'knowledge representation' (but should really be called 'belief representation', since that which is represented need only be believed; it need not be justified nor, more importantly, need it be true). In this paper, Orilia explores "[t]he design of artificial agents with … sophisticated representational and deductive capacities", such as "a solution to … intensional context problems" and proposes "an alternative to [AI researcher Kurt] Konolige's [1986] modal first-order language … [Orilia's being] based on type-free property theory" (Orilia 1994: 163). He cites several of Castañeda's papers and books, primarily on quasi-indicators and guise theory (as well as my own 1986 *Cognitive Science* paper).

C.2.2.5 Others.

Finally, in computer science and AI, there are now not only grandstudents of Castañeda such as Wiebe, but even great-grandstudents, many of whom are doing work that Castañeda would have found fascinating, and all of whom—whether they are aware of it or not—have been indirectly influenced by his point of view. (See Rapaport 1998.)

C.3 Potential, or "Ought-To-Be", Influence.

I will conclude this survey with a brief return to two areas in AI where Castañeda's influence *ought to be* more than it *is*, or perhaps I should say that some AI researchers *ought to do* more with Castañeda's theories!

C.3.1 Quasi-Indicators.

The first of these areas is quasi-indicators. It always astonishes me when I read something that is about quasi-indicators *in all but name*. I have in mind a recent article by Adam J. Grove, an AI researcher at NEC Research Institute, "Naming and Identity in Epistemic Logic," which appeared in *Artificial Intelligence* (1995) (another paper with a title that would be equally at home in a philosophy journal!). Grove 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" (Grove 1995: 314). Although he is not concerned with natural-language understanding or cognitive modeling (cf. Grove 1995: 320), he 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" (Grove 1995: 326; cf. p. 318). He introduces "a special name *I* that allows the agent to refer to himself" (Grove 1995: 319; *sic*), 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' " (Grove 1995: 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. (Grove 1995: 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 I 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.

C.3.2 Intensional Knowledge Representation.

The other area where Castañeda's theories have not been as influential as, I think, they should be is intensional knowledge representation. In an important 1991 paper in *Artificial Intelligence*, "Existence Assumptions in Knowledge Representation," Graeme Hirst, an AI researcher at the University of Toronto, argues for the importance of intensional—in particular, Meinongian—theories for computational natural-language semantics. The article is a brilliant survey of the problems and the literature but, sadly, does not mention Castañeda at all.[5]

And in another significant paper on computational semantics for natural-language processing, "Ontological Promiscuity" (1985), Jerry Hobbs, a computational linguist with the AI Center at SRI International, discusses opacity, the *de re/de dicto* distinction, and identity in intensional contexts—all without reference to Castañeda (or even, for that matter, to Meinong; curiously, Hobbs insists that a theory that admits nonexistents and impossible objects into its ontology is *Platonic*, not Meinongian).

C.4 Summary.

Before summing up, I should note that Castañeda published (to my knowledge) only one paper that talks about AI explicitly: "The Reflexivity of Self-Consciousness: Sameness/Identity, Data for Artificial Intelligence" (1989a).[6] But despite mentioning AI in the title, the only thing he has to say about it in the paper itself is this:

> … Artificial Intelligence, whether practiced with a reductionist bent of mind or not, has a vested interest in the reflexivity of self-consciousness. Clearly, the production of facsimiles of human behavior or of mental states and activities needs only the causal dependence of the mental on the physical. Self-consciousness is the apex of mentality. (Castañeda 1989a: 28–29.)

The notion of "reflexivity of self-consciousness" is explained thus:

> Our topic is the reflexivity of self-consciousness. The reflexivity in ONE referring to ONEself as *oneself* is twofold. There is the external reflexivity of ONE referring to ONEself, and the internal reflexivity of ONE referring to something, whatever it may be, as *oneself*. We must take both

[5] Hirst has reminded me that I had recommended that, since he did not have room in his paper to deal with all the "neo-Meinongian" (my term) theories of non-existents, he should probably focus on Terence Parsons's theory. *Mea culpa.*

[6] I wonder about that comma before 'Data'; i.e., perhaps 'Sameness/Identity' should be read as an adjective modifying 'Data'.

> reflexivities into account. The internal reflexivity is the peculiar core of *self*-consciousness. (Castañeda 1989a: 28.)

However, despite the paucity of Castañeda's own discussions of AI, his contributions in deontic logic and the study of quasi-indicators have directly influenced AI research in action theory and computational linguistics, his students (and theirs, in turn) have themselves made—and are making—contributions to AI, and there is, I firmly believe, great potential for more, especially the application of guise theory to issues in knowledge representation. What remains to be done—and the reason I titled this survey a "prolegomena"—is to investigate the more indirect influences in detail.

Let me close with a remark inspired by Daniel Dennett's comment on the relationship between philosophy and AI:

> Philosophers … [Daniel Dennett] said, should study AI. Should AI workers study philosophy? Yes, unless they are content to reinvent the wheel every few days. When AI reinvents a wheel, it is typically square, or at best hexagonal, and can only make a few hundred revolutions before it stops. Philosopher's wheels, on the other hand, are perfect circles, require *in principle* no lubrication, and can go in at least two directions at once. Clearly a meeting of minds is in order. (Dennett 1978: 126.)

To which I add that Castañeda's philosophical "wheels" have provided, and have continued potential to provide, important mechanisms and insights for AI workers.[7]

References

Billington, David; De Coster, Koen; & Nute, Donald E. (1990), "A Modular Translation from Defeasible Nets to Defeasible Logics," *Journal of Experimental and Theoretical Artificial Intelligence* 2: 151–177.

Bratman, Michael E. (1983), "Castañeda's Theory of Thought and Action," in James E. Tomberlin (ed.), *Agent, Language, and the Structure of the world: Essays Presented to Hector-Neri Castañeda with His Replies* (Indianapolis, IN: Hackett).

Bratman, Michael E. (1987), *Intention, Plans, and Practical Reason* (Cambridge, MA: Harvard University Press).

Bratman, Michael E. (1990), "What Is Intention?", in Philip R. Cohen, Jerry Morgan, & Martha E. Pollack (eds.), *Intentions in Communication* (Cambridge, MA: MIT Press): 15–31.

Bringsjord, Selmer (1992), "CINEWRITE: An Algorithm-Sketch for Writing Novels Cinematically, and Two Mysteries Therein," *Instructional Science* 21: 155–168;

[7]I am grateful to Andy Haas, Graeme Hirst, and Anne Reboul for comments on an earlier draft of this chapter, which was presented to the Society for Iberian and Latin American Thought at the American Philosophical Association Eastern Division meetings (New York, 1995).

reprinted in Mike Sharples (ed.), *Computers and Writing: State of the Art* (Dordrecht: Kluwer Academic Publishers, 1992).

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

Castañeda, Hector–Neri (1968), "On the Logic of Attributions of Self-Knowledge to Others," *Journal of Philosophy* 64: 439–456.

Castañeda, Hector-Neri (1970a), "On the Semantics of the Ought-to-Do," *Synthese* 21: 449–468.

Castañeda, Hector–Neri (1970b), "On the Philosophical Foundations of the Theory of Communication: Reference," *Midwest Studies in Philosophy* 2 (1977) 165–186.

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

Castañeda, Hector-Neri (1981), "The Paradoxes of Deontic Logic: The Simplest Solution to All of Them in One Fell Swoop," in Risto Hilpinen (ed.), *New Studies in Deontic Logic* (Dordrecht: D. Reidel): 37–85.

Castañeda, Hector-Neri (1989a), "The Reflexivity of Self-Consciousness: Sameness/ Identity, Data for Artificial Intelligence," *Philosophical Topics* 17: 27–58.

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

Cohen, Philip R., & Levesque, Hector J. (1990), "Persistence, Intention, and Commitment," in Philip R. Cohen, Jerry Morgan, & Martha E. Pollack (eds.), *Intentions in Communication* (Cambridge, MA: MIT Press): 33–69.

Cohen, Philip R.; Morgan, Jerry; & Pollack, Martha E. (eds.) (1990a), *Intentions in Communication* (Cambridge, MA: MIT Press).

Cohen, Philip R.; Morgan, Jerry; & Pollack, Martha E. (1990b), "Introduction," in Philip R. Cohen, Jerry Morgan, & Martha E. Pollack (eds.), *Intentions in Communication* (Cambridge, MA: MIT Press): 1–13.

Covington, Michael; Nute, Donald E.; & Vellino, André (1988), *Prolog Programming in Depth* (Glenview, IL: Scott, Foresman).

Creary, Lewis G. (1979), "Propositional Attitudes: Fregean Representation and Simulative Reasoning," *Proceedings of the 6th International Joint Conference on Artificial Intelligence (IJCAI-79, Tokyo)* (Los Altos, CA: Morgan Kaufmann), Vol. I, pp. 176–181.

Creary, Lewis G. (1983), "NFLT: A Language of Thought for Reasoning about Actions" (Stanford, CA: Stanford University Department of Computer Science Artificial Intelligence Laboratory).

Creary, Lewis G., & Pollard, Carl J. (1985), "A Computational Semantics for Natural Language," *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics (University of Chicago)* (Morristown, NJ: Association for Computational Linguistics): 172–179.

Dennett, Daniel C. (1978), "Artificial Intelligence as Philosophy and as Psychology," in Daniel C. Dennett, *Brainstorms: Philosophical Essays on Mind and Psychol-*

ogy (Montgomery, VT: Bradford Books).

Duchan, Judith F.; Bruder, Gail A.; & Hewitt, Lynne E. (eds.) (1995), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates).

Geerts, Patricia; Nute, Donald E.; & Vermeir, Dirk (1994), "Ordered Logic: Defeasible Reasoning for Multiple Perspectives," *Decision Support Systems* 11: 157–190.

Gonzalez; Morris, Robert A. McKenzie; Carreira; & Gann (forthcoming), "Model-Based, Real-Time Control of Electrical Power Systems," *IEEE Transactions on Systems, Man, and Cybernetics*.

Gorokhovski, Vincent, & Nute, Donald E. (forthcoming), "Two-Level Geological Modeling: A Computational Approach to Teaching Decision Making under Conditions of Geological Data Paucity," *Environmental and Engineering Geoscience*.

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

Grosz, Barbara J., & Kraus, Sarit (1993), "Collaborative Plans for Group Activities," *Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93; Chambéry, France)* (San Mateo, CA: Morgan Kaufmann): 367–373.

Grosz, Barbara J., & Sidner, Candace L. (1990), "Plans for Discourse," in Philip R. Cohen, Jerry Morgan, & Martha E. Pollack (eds.), *Intentions in Communication* (Cambridge, MA: MIT Press): 417–444.

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

Hilpinen, Risto (1993), "Actions in Deontic Logic," in John-Jules Ch. Meyer & Roel J. Wieringa (eds.), *Deontic Logic in Computer Science: Normative System Specification* (Chichester, UK: John Wiley & Sons): 85–100.

Hirst, Graeme (1991), "Existence Assumptions in Knowledge Representation," *Artificial Intelligence* 49: 199–242.

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.

Kaplan, David (1989), "Demonstratives," in Joseph Almog, John Perry, & Howard Wettstein (eds.), *Themes from Kaplan* (Oxford University Press): 481–563.

Karickhoff, S. W. et al. (1991), "Predicting Chemical Reactivity by Computer," *Environmental Toxicology and Chemistry* 10: 1405–1416.

Konolige, Kurt (1986), *A Deduction Model of Belief* (Los Altos, CA: Morgan Kaufmann).

Krogh, Christen (1996), "Getting Personal: Some Notes on the Relationship between Personal and Impersonal Obligation," in Carmo & Brown (eds.), *Proceedings of Deon'96* (Springer-Verlag).

Lespérance, Yves (1989), "A Formal Account of Self-Knowledge and Action," in N. S. Sridharan (ed.), *Proceedings of the 11th International Joint Conference on Artifi-*

cial 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.

Macura, Katarzyna; Macura, Robert; & Nute, Donald E. (1992), "Intelligent Tutoring System for Neuro-Oncology" (revised), *Journal of Medical Educational Technologies* (Summer): 18–23.

Macura, Katarzyna; Macura, Robert; & Nute, Donald E. (1993), "Medical Reference System for Differential Diagnosis of Brain Tumors: Motivation and Structure," *Microcomputer Applications* 12: 110–117.

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.

McCarty, L. Thorne (1983), "Permissions and Obligations," *Proceedings of the 8th International Joint Conference on Artificial Intelligence (IJCAI-83; Karlsruhe, W. Germany)* (Los Altos, CA: Morgan Kaufmann): 287–294.

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

Meyer, Bernd; Hansen, Torben; Nute, Donald E.; Albersheim, Peter; Darvill, Alan; York, William; & Sellers, Jeffrey (1991), "Identification of 1H-NMR Spectra of Complex Oligosaccharides with Artificial Neural Nets," *Science* 251: 542–544.

Meyer, John-Jules Ch., & Wieringa, Roel J. (eds.) (1993a), *Deontic Logic in Computer Science: Normative System Specification* (Chichester, UK: John Wiley & Sons).

Meyer, John-Jules Ch., & Wieringa, Roel J. (1993b), "Deontic Logic: A Concise Overview," in John-Jules Ch. Meyer & Roel J. Wieringa (eds.), *Deontic Logic in Computer Science: Normative System Specification* (Chichester, UK: John Wiley & Sons): 3–16.

Morris, Robert A., & Al-Khatib, Lina (1991), "An Interval-Based Temporal Relational Calculus for Events With Gaps," *Journal of Experimental and Theoretical Artificial Intelligence* 3: 87–107.

Morris, Robert A., & Marchetti, J. (1989), "Representing Event Identity in Semantic Analysis: A Network Approach," *Proceedings of the 2nd Florida Artificial Intelligence Research Symposium (FLAIRS-89; Orlando, FL)*: 193–196.

Morris, Robert A.; Shoaff, William D.; & Khatib, Lina (1993), "Path Consistency in a Network of Non-Convex Intervals," in Ruzena Bajcsy (ed.), *Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93; Chambéry, France)* (San Mateo, CA: Morgan Kaufmann): 655–660.

Morris, Robert A.; Shoaff, William D.; & Khatib, Lina (1996), "Domain Independent Temporal Reasoning with Recurring Events," *Computational Intelligence* 12: 450–477.

Nute, Donald E. (1988a), "Defeasible Reasoning and Decision Support Systems," *Decision Support Systems* 4: 97–110.

Nute, Donald E. (1988b), "Defeasible Reasoning: A Philosophical Analysis in Prolog," in James Fetzer (ed.), *Aspects of Artificial Intelligence* (Dordrecht: Kluwer Academic Publishers): 251–288.

Nute, Donald E. (1990a), "Defeasible Logic and the Frame Problem," in Henry Kyburg, Ronald Loui, & Greg Carlson (eds.), *Knowledge Representation and Defeasible Reasoning* (Dordrecht: Kluwer Academic Publishers): 3–21.

Nute, Donald E. (1990b), "SCHOLAR: A Scholarship Identification System in Prolog," in Rao Aluri & Donald E. Riggs (eds.), *Expert Systems in Libraries* (Norwood, NJ: Ablex): 98–108.

Nute, Donald E. (1992), "Basic Defeasible Logic," in L. Fariñas del Cerro & M. Penttonen (eds.), *Intensional Logics for Programming* (Oxford: Oxford University Press): 125–154.

Nute, Donald E. (1993), "Inference, Rules, and Instrumentalism," *International Journal of Expert Systems Research and Applications* 5: 267–274.

Nute, Donald E. (1994), "Defeasible Logic," in Dov M. Gabbay, C. J. Hogger, & J. A. Robinson, eds., J. Siekmann, volume co-ordinator, *Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol. 3, *Nonmonotonic Reasoning and Uncertain Reasoning* (Oxford: Clarendon Press): 353–395.

Nute, Donald E. (ed.) (forthcoming), *Defeasible Deontic Logic: Essays in Nonmonotonic Normative Reasoning* (Dordrecht: Kluwer Academic Publishers).

Nute, Donald E.; Mann, Robert; & Brewer, Betty (1990), "Controlling Expert System Recommendations with Defeasible Logic," *Decision Support Systems* 6: 153–164.

Nute, Donald E.; Rauscher, H. Michael; Perala, Donald A.; Zhu, Guojun; Chang, Yousong; & Host, George M. (1995), "A Toolkit Approach to Developing Forest Management Advisory Systems in Prolog," *AI Applications* 9.3: 39–58.

Orilia, Francesco (1992), "Intelligenza artificiale e proprietà mentali," *Nuova civiltà delle macchine* 10.2 (38), pp. 44–63

Orilia, Francesco (1994a), "Artificial Agents and Mental Properties," in Roberto Casati, Barry Smith, & Graeme White (eds.), *Philosophy and the Cognitive Sciences*, (Vienna: Hölder-Pichler-Tempsky).

Orilia, Francesco (1994b), "Belief Representation in a Deductivist Doxastic Logic," *Minds and Machines* 4: 163–203.

Orilia, Francesco (1994c), "Side-Effects and Cohen's and Levesque's Theory of Intention as a Persistent Goal", *From the Logical Point of View* 3: 14–31.

Orilia, Francesco (1995), "Knowledge Representation, Exemplification, and the Gupta–Belnap Theory of Circular Definitions," in Gori & Soda (eds.), *Topics in Artificial Intelligence: Proceedings of the 4th Italian Conference of the Italian Association*

for AI (Berlin: Springer-Verlag).

Perry, John (1979), "The Problem of the Essential Indexical," *Noûs* 13: 3–21.

Pollack, Martha E. (1991), "Overloading Intentions for Efficient Practical Reasoning," *Noûs* 25: 513–536.

Potter, W. D.; Tidrick, T. H.; Hansen, Torben M.; Hill, L. W.; Nute, D. E.; Fried, L. M.; & Leigh, T. W. (1990), "SLING: A Knowledge-Based Product Selector," *Expert Systems with Applications* 1: 161–169.

Rapaport, William J. (1976), *Intentionality and the Structure of Existence*, Ph.D. dissertation (Bloomington: Indiana University Department of Philosophy).

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

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

Rapaport, William J. (compiler) (1998), "Academic Family Tree of Hector-Neri Castañeda", in Francesco Orilia & William J. Rapaport (eds.), *Thought, Language, and Ontology: Essays in Memory of Hector-Neri Castañeda* (Dordrecht: Kluwer Academic Publishers): 369–374.

Rapaport, William J. (forthcoming), "Meinongian Semantics and Artificial Intelligence," in Peter Simons (ed.), *Essays on Meinong* (Amsterdam: Rodopi Editions).

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. (1995), "Cognition and Fiction," in Judith Felson Duchan, Gail A. Bruder, & Lynne Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 107–128.

Rapaport, William J.; Shapiro, Stuart C.; & Wiebe, Janyce M. (1997), "Quasi-Indexicals and Knowledge Reports," *Cognitive Science* 21: 63–107; reprinted in Francesco Orilia & William J. Rapaport (eds.), *Thought, Language, and Ontology: Essays in Memory of Hector-Neri Castañeda* ((Dordrecht: Kluwer Academic Publishers, 1998): 235–294.

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

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

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

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; shorter version appeared in *Proceedings of the 5th National Conference on Artificial Intelligence (AAAI–86, Philadelphia)* (Los Altos, CA: Morgan Kaufmann): 278–283; a revised version of the shorter version appears as "A Fully Intensional Propositional Semantic Network," in Leslie Burkholder (ed.), *Philosophy and the Computer* (Boulder, CO: Westview Press, 1992): 75–91.

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; reprinted in Fritz 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 Felson Duchan, Gail A. Bruder, & Lynne Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 79–105.

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.