Steps Toward Formalizing Context

Varol Akman[*] Mehmet Surav[†]

I wish honorable gentlemen would have the fairness to give the entire context of what I did say, and not pick out detached words. (Cobden, Speeches *46, 1849, quoted in the* OED*)*

The importance of contextual reasoning is emphasized by various researchers in AI. (A partial list includes John McCarthy and his group; R. V. Guha; Yoav Shoham; Giuseppe Attardi and Maria Simi; Fausto Giunchiglia and his group.) Here, we will survey the problem of formalizing context and explore what is needed for an acceptable account of this abstract notion.

The issue of context arises in assorted areas of AI, including knowledge representation, natural language processing, and intelligent information retrieval. Although the word *context* is frequently employed in descriptions, explanations, and analyses of computer programs in these areas, its meaning is frequently left to the reader's understanding, i.e., it is used in an implicit and intuitive manner.[1]

An example of how contexts may help in AI is found in McCarthy's (constructive) criticism [42] of MYCIN [57], a program for advising physicians on treating bacterial infections of the blood and meningitis. When MYCIN is told that the patient has *Chlorae Vibrio* in his intestines, it would immediately recommend two weeks of tetracycline treatment and nothing else. While this would indeed do away with the bacteria, the patient would perish long before that due to diarrhea. A "contextual" version of MYCIN should know about the context of a treatment and would realize that any prescription must be made in the light of the fact that there is alarming dehydration. Thus, in the contextual MYCIN, the circumstances surrounding a patient would have to be made explicit using a formal approach and would be used as such by the program.

[*]Department of Computer Engineering and Information Science, Bilkent University, Bilkent, Ankara 06533, Turkey. E-mail: `akman@bilkent.edu.tr`. Phone/fax: +90 312 266–4133/–4126. Please address all future correspondence to this author.

[†]Engineering and Business Systems, Inc., 2 Enterprise Drive, Shelton, Connecticut 06484, USA. E-mail: `surav@dacka.ebs.com`.

[1]In an early influential paper, Clark and Carlson [19] state that context has become a favorite word in the vocabulary of cognitive psychologists and that it has appeared in the titles of a vast number of articles. They then complain that the the denotation of the word has become murkier as its uses have been extended in many directions and deliver the now widespread opinion that context has become some sort of "conceptual garbage can."

The main motivation for studying formal contexts is to resolve the *problem of generality* in AI, as introduced by McCarthy [43]. McCarthy believes that AI programs suffer from a lack of generality. A seemingly minor addition (as in the MYCIN example) to the particular, predetermined possibilities that a program is required to handle often necessitates a partial redesign and rewrite of the program. Explicitly represented contexts would help because a program would then make its assertion *about* a certain context.

A more general objection to the implicit representation of context—unless the representation is part of a program which is not "mission-critical" (in a broad sense of the term)—can be given as follows. Assume that we write longer (more involved in terms of complexity) or shorter (simpler in terms of complexity) axioms depending on which implicit context we are in. The problem is that, long axioms are often longer than is convenient in daily situations. Thus, we find it handy to utter "This is clever," leaving any explanation as to whether we are talking about a horse or a mathematical argument to the context of talk. On the other hand, shorter axioms may invite just the opposite of a *principle of charity* from an adversary. To quote McCarthy [43, p. 1034]:

> Consider axiomatizing *on* so as to draw appropriate consequences from the information expressed in the sentence, "The book is on the table." The [adversary] may propose to haggle about the precise meaning of *on*, inventing difficulties about what can be between the book and the table, or about how much gravity there has to be in a spacecraft in order to use the word *on* and whether centrifugal force counts.

While our aim in this paper is to offer a review of recent formalizations of context—those that can be used for automated reasoning—we will first identify the role of context in various fields of AI. Next section does this and may be regarded as the fundamental motivational basis for our study. Following that, we will consider some (logic-based) attempts towards formalizing context. In that section, the focus of our discussion will be McCarthy's proposal [45], which, in our view, is the groundwork for all other logicist formalizations.

The approach which we pursue in our own line of research is inspired by Situation Theory (cf. Barwise and Perry [12] and especially Devlin [20]) and is detailed in Barwise [8]. The essence of Barwise's proposal will be reviewed in a separate section. Objectivity and prudence dictate that we do not review our own work here. Therefore, we refer the interested reader to two recent papers which detail our standpoint [59, 1].[2]

Some Useful Definitions

According to the *Oxford English Dictionary* this term usually has two primary meanings[3]: (i) the words around a word, phrase, statement, etc. often used to

[2] The two workshop proceedings in which our papers appear may also be an excellent starting point to get a taste of current research in contextual reasoning.

[3] *OED* also exemplifies several other meanings most of which are not currently used. Incidentally, our opening quotation of this review also comes from this dictionary.

help explain (fix) the meaning; (ii) the general conditions (circumstances) in which an event, action, etc. takes place. Clearly, the first meaning is closely related to linguistic meaning and linguists' use of the term, whereas the second—more general—meaning is the one which is closer to a desirable account of context in AI.[4] In *The Dictionary of Philosophy* [4], the same term is defined, reflecting the latter desideratum more satisfactorily, as follows:

> **context** (L. *contexere*, "to weave together." from *con*, "with," and *texere*, "to weave"): The sum total of meanings (associations, ideas, assumptions, preconceptions, etc.) that (a) are intimately related to a thing, (b) provide the origins for, and (c) influence our attitudes, perspectives, judgments, and knowledge of that thing.

Similarly, in another dictionary (*Collins Cobuild English Language Dictionary*), the prevalent meanings of the term include the following:

- The *context* of something consists of the ideas, situations, events, or information that relate to it and make it possible to understand it fully.
- If something is seen *in context* or if it is put *into context*, it is considered with all the factors that are related to it rather than just being considered on its own, so that it can be properly understood.
- If a remark, statement, etc. is taken or quoted *out of context*, it is only considered on its own and the circumstances in which it was said are ignored. It, therefore, seems to mean something different from the meaning that was intended.

The Role of Context

Context in Natural Language

Context is a crucial factor in communication.[5] Ordinary observation proves its importance in assorted ways. Just consider the confusion which results from lack of contextual information when, for example, you join a scheduled meeting half an hour late. Without the clues of the original context, you might find it hard to make sense of the ongoing discussion. In any case, those discussants would realize that they cannot assume a lot about the background knowledge you possess and would normally give you a quick rundown of the conversations so far. This is essentially the

[4] This is in the spirit of the following observation of McCarthy [44, p. 180]: "Almost all previous discussion of context has been in connection with natural language, and the present paper relies heavily on examples from natural language. However, I believe the main AI uses of formalized context will not be in connection with communication but in connection with reasoning about the effects of actions directed to achieving goals. It's just that natural language examples come to mind more readily."

[5] Linguists have talked about context for over a century and have many interesting results. Much, much more would need to be said about context in linguistics but since we cannot hope to do justice to such a great body of work in this review, we will be content with a brief, superficial appraisal.

view of Clark and Carlson [19] who regard context as information that is available to a person for interaction with a particular process on a given occasion. Their *intrinsic context* is an attempt to capture the information available to a process that is potentially necessary for it to succeed. The intrinsic context for grasping what a speaker means on some occasion is the (limited) totality of the knowledge, beliefs, and suppositions that are shared by the speaker and the listener (a.k.a. the *common ground*).

Leech [38, p. 66] gives another particularly attractive quasi-definition as follows:

> [W]e may say that the specification of context (whether linguistic or non-linguistic) has the effect of narrowing down the communicative possibilities of the message as it exists in abstraction from context.

Thus, context is seen as having a so-called disambiguating function (among others). To quote Leech [38, p. 67] once again "The effect of context is to attach a certain *probability* to each sense (the complete ruling-out of a sense being the limiting case of *nil* probability)." Consider the following simple (possibly trivial for human beings) segment of conversation [9]:

A (a woman, talking to B): *I am a philosopher.*

B (talking to C and referring to A): *She is a philosopher.*

C (talking to A): *So, you are a philosopher.*

Context eliminates certain ambiguities or multiple meanings in the message. In the above segment, one of the very first context dependent things is the word "philosopher." The meaning of this word is determined using the context of conversation. Although the above excerpt is not sufficient to carry the proper connotation of this word, our common understanding selects an appropriate meaning from a set of possible meanings.[6]

In the above example, the indexicals (e.g., I, you) can be bound to appropriate persons only by the help of context. For example, the sentences uttered by A and B have the same content, and we can only say this using some circumstantial information and conventions about the conversation. This circumstantial information might be formalized via context. To quote Recanati [51, p. 235], "[T]he meaning of a word like 'I' is a function that takes us from a context of utterance to the semantic value of the word in that context, which semantic value (the reference of 'I') is what the word contributes to the proposition expressed by the utterance." This is the view made popular by Kaplan's seminal work on the logic of demonstratives [37].

Another function of context arises when we deal with quantifiers in logic or natural language semantics. The range and interpretation of quantifiers depend on the context. For example, the quantifier "all" usually does not apply to all objects,

[6] According to the *Merriam-Webster Dictionary*, this word may stand for any of the following: a reflective thinker (scholar); a student of or specialist in philosophy; one whose philosophical perspective enables him to meet trouble calmly. Most probably, it is the second meaning which is ordinarily evoked in the mind of the reader.

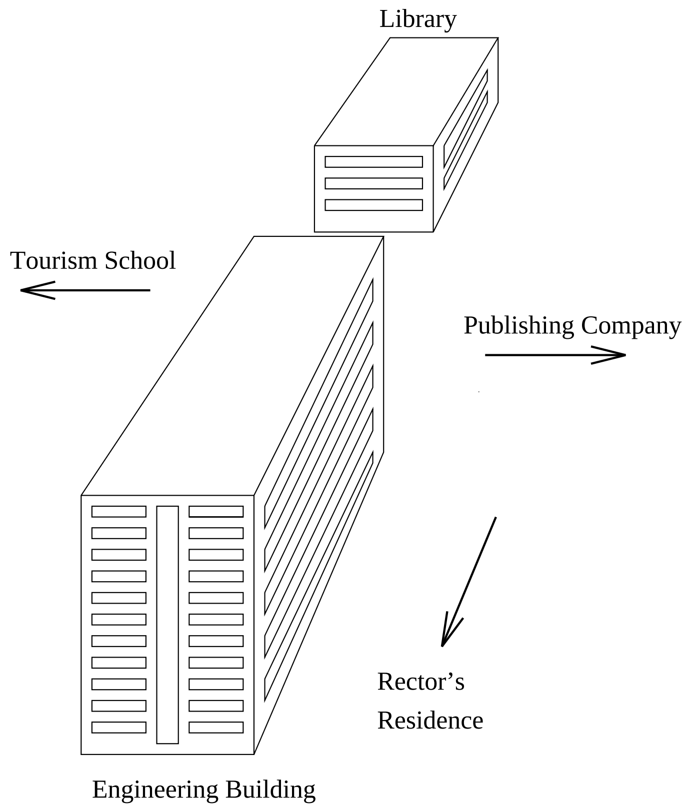

Figure 1: A partial view of the Bilkent Campus.

but only to those of a particular kind in a particular domain, determined by the contextual factors. Another example might be the interpretation of the meaning of "many." In an automobile factory, 10 automobiles might not qualify as many, but if a person owns 10 automobiles it counts as many. Clearly, even the last interpretation is context dependent. One might propose that "many" can only be interpreted as a ratio. But this, too, has a contextual dependency on the ratio. In a class of students, half of the students cannot be considered as many to cancel a midterm exam, but surely must be regarded as many in an influenza epidemic.

Context might be used to fill the missing parameters in natural language utterances. Consider an utterance of the sentence

Carl Lewis is running.

Here, the time and the place of the running action are determined by the context. For example, if we are watching a competing Lewis on TV at the 1992 Barcelona Olympic Games, then the time and the place of the utterance are different from what we would get if we are watching him practice from our window.

Some relations stated in natural language necessarily need a context for disambiguation. Consider an utterance of the sentence

The Engineering Building is to the left of the Library.

In the context of the Bilkent Campus (see Figure 1), if we are viewing the buildings from the Publishing Company, the utterance is true, but if we are in the Tourism School, the utterance is false. More interestingly, if we are looking from the Rector's Residence, this utterance must be considered neither true nor false: the Library is behind the Engineering Building.

In the light of these examples, it is seen that for natural languages a *fleshing-out* strategy, i.e., converting everything into decontextualized eternal sentences [50], cannot be employed since we do not always have full and precise information about the relevant circumstances.

Several studies in computational linguistics focused on the semantics of coherent multi-sentence discourse (or text).[7] The essential idea is that in discourse each new sentence s should be interpreted in the context provided by the sentences preceding it. As a result of this interpretation, the context is enriched with the contribution made by s. (For example, an important aspect of this enrichment is that elements are introduced that can serve as antecedents to anaphoric expressions following s.) With its emphasis on representing and interpreting discourse in context, Discourse Representation Theory [36, 61] has influenced much subsequent work in computational linguistics.

To interpret extended discourse, some other researchers regard discourse as a hierarchically organized set of segments. The expectation is that each segment displays some sort of local coherence, i.e., it can be viewed as stressing the same point or describing the same state of affairs. Grosz and Sidner [29] outline a particularly illuminating model of this segmentation process, complete with an intentional constitution of discourse available in the segments.

Using general world knowledge (e.g., knowledge about causality and everyday reasoning) is also fundamental for discerning much discourse. An early work that used such knowledge for computer understanding of natural language is due to Schank and Rieger [54]. More recent technical contributions by Charniak [18] and Hobbs et al. [35] may be seen as two formal attempts in this regard, viz. to generate expectations that are matched into plausible interpretations of the discourse, and to construct an (abduction-based) argument that explains why the current sentence is true, respectively.

Finally, it is worth noting that context has long been a key issue in social studies of language, viz. how human beings employ language to build the social and cultural organizations that they inhabit. Lyons thinks that this is natural, affirming that "in the construction of a satisfactory theory of context, the linguist's account of the interpretation of utterances must of necessity draw upon, and will in turn contribute to, the theories and findings of social sciences in general: notably of psychology, anthropology and sociology" [40, p. 292]. For a recent appreciation of work of this

[7] Establishing coherence is an integral part of the interpretation task. Allen [3, p. 465] explains this nicely: "A discourse is coherent if you can easily determine how the sentences in the discourse are related to each other. A discourse consisting of unrelated sentences would be very unnatural. To understand a discourse, you must identify how each sentence relates to the others and to the discourse as a whole. It is this assumption of coherence that drives the interpretation process."

sort, the reader is referred to Goodwin and Duranti [28] (and other articles in the same volume). They consider context a key concept in ethnographically oriented studies of language use and claim that this notion "stands at the cutting edge of much contemporary research into the relationship between language, culture, and social organization, as well as into the study of how language is structured in the way it is" [28, p. 32].

Context in Categorization

Categorization is one of the basic mental processes in cognition [53]. We, as human beings, can categorize various types of objects, events, and states of affairs, and our categorizations depend on the circumstance and perspective. Consider the following scenario:

> In Springfield (the home-town of Bart Simpson), there are three barbers working for money, and a man who does not work for money (since he has another job) but serves the community by shaving senior citizens on Sundays. If we look at the situation from a commonsense perspective, there are four barbers in town, but from say, the mayor's point of view, there are only three (licensed, tax-paying, etc.) barbers.

Here, it is clear that context (or perspective) plays an important part in the correct classification.

Barwise and Seligman [13] use *natural regularities* to study the role of context in categorization. An example regularity from Seligman [55] is *Swans are white.* This is a typical natural regularity in the sense that it is both reliable and fallible. Natural regularities are reliable since they are needed to explain successful representation, knowledge, truth, and correct reference. And they are fallible since they are needed to account for misinterpretation, error, false statements, and defeasible reference. Swans are in general white, thus the regularity is reliable and explains a fact. There might be exceptions like the Australian swans—they are usually black—but this does not mean that the regularity does not hold. Here, the fundamental problem with isolating the essential properties of a regularity is that any statement of them depends on some context of evaluation, i.e., we should evaluate the above regularity for say, the European swans.

There is a correlation between nonmonotonic reasoning and the role of context dependent factors in natural regularities. Although natural regularities are usually considered in philosophical discussions, they intuitively correspond to material implication in logic, and the effect of contextual factors is similar to the effect of nonmonotonicity. The difference between the philosophical and the logical approaches is in the way that these two disciplines differ. In logic, implication and nonmonotonicity are usually studied in a syntactic fashion, and the reasons behind the abnormalities are usually left out of the scope of discussion.

If we could completely describe all the contextual factors, then the problem would go away and we would not require extra machinery. However, we must always include a "so forth" part to cover the unexpected contextual factors; in many cases, it is simply impossible to state all of the relevant ones [60]. Still, we must somehow be able to deal with them. This is why the introduction of the notion of context, and using this in categorization should be useful.

Context in Intelligent Information Retrieval

A formal notion of context might be useful in IR since it can increase the performance by providing a framework for well-defined queries and intelligent text matching. Given the explicit context, a query may be better described and thus the recall and precision may be enhanced. In this sense, we find the work of Hearst [34] useful, for she emphasizes the importance of context in full-text information access.

Traditional methods of IR use statistical methods to find the similarities between the documents and relevance of the documents to the query. In this respect, a formal context means that the query will be better described since it will contain more information than a few keywords in the search. Inclusion of the context of the query also allows us to run more sophisticated methods in measuring the relevance.

There are various syntactical approaches to measure the relevance of a term to a document. Until recently, the only respectable methods were the statistical methods which are based on the frequency of occurrence. Lately, psychological, epistemic, and semantical considerations are beginning to flourish [21]. For example, Park [49] studies the contributions of relevance to improving information retrieval in public libraries. According to her, the search criteria for any query should be set according to the users' criteria of relevance. Since different users exhibit different relevance criteria, the query formation is a dynamic task.

The essential work on relevance is due to Sperber and Wilson [58], in which relevance is mainly considered to be the psychological relevance of a proposition to a context. Their assumption is that people have intuitions of relevance, i.e., they can consistently distinguish relevant from irrelevant information. However, these intuitions are not very easy to elicit or use as evidence, since the ordinary language notion of relevance comes along with a fuzzy (variable) meaning. Moreover, intuitions of relevance are relative to contexts, and there is no way of controlling exactly which context someone will have in mind at a given moment. Despite these difficulties, Sperber and Wilson intend to invoke intuitions of relevance. According to them, a proposition is relevant to a context if it interacts in a certain way with the (context's) existing assumptions about the world, i.e., if it has some contextual effects. These contextual effects include: (i) *Contextual implication*: A new assumption can be used together with the existing rules in the context to generate new assumptions; (ii) *Strengthening*: A new assumption can strengthen some of the existing assumptions; and (iii) *Contradicting or eliminating*: A new assumption may change or eliminate some of the existing assumptions of the context.

Sperber and Wilson talk about "degrees" of relevance. Clearly, one piece of information may be more relevant to a particular context than another. To compare the relevance of pieces of information, they consider the mental processing effort, i.e., the length of the chain of reasoning and the amount of encyclopedic information involved, and so on. Finally, they propose their celebrated *Relevance Maxim* [58] which has two parts:

1. An assumption is relevant in a context to the extent that its contextual effects in this context are large.
2. An assumption is irrelevant in a context to the extent that the effort required to process it in this context is large.

Harter [33] uses the theoretical framework of Sperber and Wilson to interpret psychological relevance in relation to IR. According to him, reading a new bibliographic citation (the setting here is that of a user, accepting or rejecting a bibliographic document retrieved by a library information system) can cause a user to create a new context. A set of cognitive changes take place in that context and the citation and the context influence each other to give rise to new ideas. In other words, a retrieved citation (viewed as a psychological stimulus) is relevant to a user if it leads to cognitive changes in that user.

Using knowledge about data to integrate disparate sources may be considered as a more sophisticated extension of IR and is also being actively researched [27]. The motivation is that while networking technologies make physical connectivity (and hence access to assorted data) feasible, much of this data is not perceived as meaningful due to a lack of information regarding the context of the data. This research aims at the development of a formal theory of *context interchange* [52] using (i) context definitions (e.g., defining the semantics, organization, and content of data) and (ii) context characteristics (e.g., data quality, security, etc.), and thus suggests a solution to the problem of semantic interoperation between (semantically) heterogeneous environments.

Context in Knowledge Representation and Reasoning

When we state something we do so in a context. For example, 37 degrees centigrade is "high" in the context of a weather report, but "normal" in the context of a medical diagnosis. In the context of Newtonian mechanics time is ethereal, but in the context of general relativity, this is hardly the case. The examples can be continued. The main point is that if we are to reason in a common sense way, we have to use certain contexts.

The importance of the notion of context was realized by philosophers[8] for centuries. Early on, philosophers recognized that a causal connection between two events holds only relative to a certain background, thus only in certain contexts.

[8]Cf. Wittgenstein [62, p. 166]: " 'He measured him with a hostile glance and said ....' The reader of the narrative understands this; he has no doubt in his mind. Now you say: 'Very well, he supplies the meaning, he guesses it.'—Generally speaking: no. Generally speaking he supplies nothing, guesses nothing.—But it is also possible that the hostile glance and the words later prove to have been pretense, or that the reader is kept in doubt whether they are so or not, and so that he really does guess at a possible interpretation.—But then the main thing he guesses at is a context. He says to himself for example: The two men who are here so hostile to one another are in reality friends, etc. etc."

McCarthy was the first researcher to realize that the introduction of a formal notion of context is required for generality in AI [43].

According to McCarthy, there is simply no most general context in which all the stated axioms always hold and everything is meaningful. When one writes an axiom, it holds in a certain context, and one can always present another (more general) context in which the axiom fails. McCarthy formalizes relativized-truth-within-a-context via a special predicate $\texttt{holds}(p,c)$. This states that proposition p holds in context c.[9]

If we compare the two approaches, namely, holds and adding a context parameter to each function and predicate, we must prefer holds, since it allows us to use the notion of context uniformly as the other objects.[10] A problem with this approach is that, if we are to live in a first order world (the world of FOL), we have to reify p in $\texttt{holds}(p,c)$. Alternative (modal) approaches to reifying assertions are investigated in [56, 16, 48].

Among the advantages gained as a result of the use of contexts are the following:[11]

- *Economy of representation*: Different contexts can circumscribe (in a nontechnical sense) the parts of the knowledge base that are accessible in different ways, thus allowing the representation of many knowledge bases in a single structure.

- *Efficiency of reasoning*: By factoring out a possibly very large knowledge base, contexts may permit more competent reasoning about the real, intended scope.

- *Allowing inconsistent knowledge bases*: The knowledge base might be partitioned according to the context of its use. In this way, we can accommodate contradicting information in the same knowledge base as long as we treat such information carefully.

- *Resolving lexical ambiguity*: By using context, the task of choosing the right interpretation of lexical ambiguity is made easier.[12] However, there are arguments that while a context formalism can represent lexical ambiguity, additional knowledge is needed to perform the resolution [15].

- *Flexible entailment*: Context might effect the entailment relation. For example, in a particular context, entailment might warrant a closed world assump-

[9] We will see that in McCarthy's newer work [45] $\texttt{holds}(p,c)$ is renamed $\texttt{ist}(c,p)$.

[10] This is desirable because "[f]ormalizing commonsense reasoning needs contexts as objects, in order to match human ability to consider context explicitly" [44, p. 180]. Furthermore, adding a context parameter to say, each predicate would be unnatural. Consider the following example about the predicate *at*, adapted from [47]. In one context, *at* may be taken in the sense of being regularly at a place; in another, it may mean physical presence at a certain instant. Programs based on formalized contexts would extract the appropriate meaning automatically. Otherwise, a need arises to use two different predicates, e.g., at_1 and at_2.

[11] In identifying the given computational advantages of the notion, we have heavily depended on Shoham [56, pp. 395–396].

[12] An good example due to Shoham [56] is the use of the word "glasses": the appropriate meaning would be different in the context of a wine-and-cheese party and in the context of a visit to an ophthalmologist.

tion whereas in some other context, this assumption needs to be dropped (the classical case).

Being the largest commonsense knowledge building attempt, CYC [32, 39] has crucial pointers on reasoning with an explicit notion of context [30]. Some aspects of the representation of knowledge which are influenced by contextual factors include

- *Language*: The language (i.e., the predicates, functions, and categories) used for representation should be chosen to be appropriate for their intended domain. For example, MYCIN and ONCOCIN—two renowned medical diagnosis programs—overlap significantly in their domains; however ONCOCIN has some concept of time whereas MYCIN does not.

- *Granularity and accuracy*: As with the vocabulary, the application area, thus context, determine the granularity and accuracy of the theory.

- *Assumptions*: The assumptions that a given task permits often lead to a simplification of the vocabulary. If we try to continue this simplification for large domains, at one point the assumptions become unstable. Thus, either we should use a highly expressive vocabulary or distribute the assumptions to different tasks.

CYC researchers identify two approaches to building large commonsense knowledge bases, and reasoning with them.

The straightforward way that a knowledge base builder might choose would be the introduction of an extremely expressive and powerful vocabulary. This approach increases the complexity of the problem, since using such a vocabulary causes difficulties in truth maintenance, and produces large search spaces.

The second way, which is Guha's way [30], is to make the context dependence of a theory explicit. In this approach, assertions (axioms, statements) are not universally true; they are only true in a context. An assertion in one context might be available for use in a different context by performing a relative decontextualization.

In CYC, the uses of context include the following:

- *A general theory of some topic*: A theory of mechanics, a theory of weather in Alabama, a theory of what to look for when buying dresses, etc. Such contexts are called *micro-theories* [30]. Different micro-theories make different assumptions and simplifications about the world. For any topic, there might be different micro-theories of that topic, at varying levels of detail.

- *A basis for problem solving*: For some difficult problems, we can form a particular context. We collect all related assumptions, rules, etc. in that context (called the *Problem Solving Context*, PSC, in CYC [32]), and can process a group of related queries in relatively small search space. Such contexts must be created dynamically and be disposed of afterwards.

- *Context dependent representation of utterances*: Naturally, we can use anaphoric and indefinite statements without completely decontextualizing them. For example, the words "the person" might be used in a discourse without identifying him/her exactly.

Formalizations in Logic

The notion of context was first introduced to AI in a logicist framework by McCarthy in 1986 in his Turing Award paper, later published as [43].[13] McCarthy published his recent ideas on context in [45, 47, 46]. Other notable works on formalizing context are due to Guha [30], S. Buvač et al. [17, 16], F. Giunchiglia et al. [25, 23], Attardi and Simi [7, 6], and Shoham [56].

We have reviewed McCarthy's early ideas [43] in the previous section. In this section, we will evaluate the other logicist formalizations, starting with McCarthy's more recent proposal.

McCarthy on Contexts

In [45], McCarthy states three reasons for introducing the formal notion of context.[14]

First, the use of context allows simple axiomatizations. He exemplifies this by stating that axioms for static blocks world situations can be *lifted*[15] to more general contexts—to contexts in which the situation changes.

Second, contexts allow us to use a specific vocabulary of and information about a circumstance. An example of this might be the context of a (coded) conversation in which particular terms have particular meanings that they would not have in the daily language in general.[16]

McCarthy's third goal is to propose a mechanism by which we can build AI systems which are never permanently stuck with the concepts they use at a given time because they can always *transcend* the context they are in. This brings about two problems:

- *When to transcend a context?* Either the system must be smart enough to do so or we must instruct it when to transcend one or more levels up.

- *Where to transcend?* This can be answered if we are prepared to accept that formulas are always considered to be asserted within a context.

[13] Contexts were not mentioned in McCarthy's Turing award talk in 1971.

[14] A newer and expanded treatment is available as [47]. The most recent reference is [46] although this is a very brief introduction.

[15] The word *lifting* is used frequently in this paper and will be discussed later.

[16] A more concrete use (from Computer Science) can be identified if we form an analogy with programming language and database concepts. McCarthy's approach corresponds to the use of *local variables* and *local functions* in programming languages, and *views* in database systems. In each case, the meaning of a term depends upon the context in which it is used.

The basic relation relating contexts and propositions is $\mathtt{ist}(c,p)$. It asserts that proposition p is true in context c. Then the main formulas are sentences of the form

$$c' : \mathtt{ist}(c,p)$$

In other words, p is true in context c, and this itself is asserted in an outer context c'.

To give an example of the use of ist,

$$c_0 : \mathtt{ist}(\textit{context-of}(\text{“Sherlock Holmes stories”}), \text{“Holmes is a detective”})$$

asserts that it is true in the context of Sherlock Holmes stories that Holmes is a detective. Here, c_0 is considered to be the *outer context*. On the other hand, in the context *context-of*("Sherlock Holmes stories"), Holmes's mother's maiden name does not have a value.

Two key properties of context are the following:

1. *Contexts are abstract objects*: Some contexts will be *rich* objects[17] just like the situations in Situation Calculus. Some contexts will not be as rich and might be fully described, e.g., simple micro-theories [30].

2. *Contexts are first-class citizens*: We can use contexts in our formulas in the same way we use other objects.

Relations and Functions Involving Contexts

There are some relations working between contexts. The most notable one is $\preceq$. This defines a partial ordering over contexts; between two contexts, we might consider a *more general than* relation ($c_1 \preceq c_2$) meaning that the second context contains all the information of the first context and probably more. Using $\preceq$, we can lift a fact from a context to one of its super-contexts using the following nonmonotonic rule:

$$\forall c_1\, \forall c_2\, \forall p\ \ (c_1 \preceq c_2) \wedge \mathtt{ist}(c_1,p) \wedge \neg ab1(c_1,c_2,p) \rightarrow \mathtt{ist}(c_2,p)$$

Here, c_2 is a super-context of c_1, p is a predicate of c_1, $ab1$ is an abnormality predicate, and $\neg ab1(c_1,c_2,p)$ is used to support nonmonotonicity. Analogously, we can state a similar lifting rule between a context and one of its sub-contexts:

$$\forall c_1\, \forall c_2\, \forall p\ \ (c_1 \preceq c_2) \wedge \mathtt{ist}(c_2,p) \wedge \neg ab2(c_1,c_2,p) \rightarrow \mathtt{ist}(c_1,p)$$

The difference between the abnormality relations is crucial: $ab1$ represents the abnormality in generalizing to a super-context, whereas $ab2$ corresponds to the abnormality in specializing to a sub-context.

Here are some examples of functions on contexts which we might want to define.

$value(c,t)$ is a function which returns the value of term t in context c:

$$value(\textit{context-of}(\text{“Sherlock Holmes stories”}), \text{“number of wives of Holmes”}) = 0$$

[17] A rich object cannot be defined or completely described. A system may be given facts about a rich object but never the complete description.

This states that Holmes has no wife in the context of Sherlock Holmes stories.

specialize-time(t, c) is a context related to c in which the time is specialized to the value t:

$$c_0 : \texttt{ist}(\textit{specialize-time}(t, c), \textit{at}(\textit{JMC}, \textit{Stanford}))$$

states that at time t in context c, John McCarthy is at Stanford University. Instead of specializing on time, we can also specialize on location, speaker, situation, subject matter, and so on.

The formal theory of context can be used to model inference in the style of deduction. Thus, $assuming(p, c)$ is another context like context c in which predicate p is assumed. Using this function, we might dynamically create a context containing the axioms that we desire. The new context validates the following rules [47]:

- *Importation*: This is the rule $c : p \rightarrow q \vdash assuming(c, p) : q$.
- *Discharge*: This is the rule $assuming(c, p) : q \vdash c : p \rightarrow q$.

When we take contexts in this *natural deduction* sense (as McCarthy suggested [43]), the operations of *entering* and *leaving* a context might be useful and shorten the proofs involving contexts. In this case, $\texttt{ist}(c, p)$ will be analogous to $c \rightarrow p$, and the operation of entering c can be taken as $assuming(p, c)$. Then, entering c and inferring p will be equivalent to $\texttt{ist}(c, p)$ in the outer context.

Lifting

Here are some of the things we can do with lifting:

- *Verbatim transfer of a formula*: If two contexts are using the same terminology for a concept in an axiom, this is a natural choice. For example, the following lifting rule states that we can use the axioms related to the *on*(x, y) of *above-theory* context in *general-blocks-world* context without any change:

  $$c_0 : \forall x \forall y\ \texttt{ist}(\textit{above-theory}, \textit{on}(x, y)) \rightarrow \texttt{ist}(\textit{general-blocks-world}, \textit{on}(x, y))$$

- *Change the arity of a predicate*: In different contexts, the same predicate might take a different number of arguments. McCarthy's example for this is *on* which takes two arguments in *above-theory* context, and three arguments in a context c in which *on* has a third argument denoting the situation. The lifting rule is

  $$c_0 : \forall x \forall y \forall s\ \texttt{ist}(\textit{above-theory}, \textit{on}(x, y)) \rightarrow \texttt{ist}(\textit{context-of(s)}, \textit{on}(x, y, s))$$

  where *context-of* is a function returning the context associated with the situation s in which the usual *above-theory* axioms hold.

- *Change the name of a predicate*: Similar to the case with arities, we can change the name of a predicate via lifting rules. For example, we can translate *on* to *üzerinde*, when we move from *above-theory* to *turkish-above-theory*:

  $$c_0 : \forall x \forall y\ \texttt{ist}(\textit{above-theory}, \textit{on}(x, y)) \rightarrow \texttt{ist}(\textit{turkish-above-theory}, \textit{üzerinde}(x, y))$$

Other Issues

Relative decontextualization is a way proposed by McCarthy to do the work of "eternal" sentences—the mythical class embracing those sentences which express the same proposition no matter in which world the utterance takes place [50]. (This obviously assumes that the world in question is linguistically similar to ours.) McCarthy feels strongly that eternal sentences do not exist. His proposed mechanism depends on the premise that when several contexts occur in a discussion, there is a common context above all of them into which all terms and predicates can be lifted. (However, the outermost context does not exist.) Sentences in this context are relatively eternal. A similar idea is used in the PSCs of CYC.

Another place where context might be useful is the representation of *mental states* [45]. McCarthy proposes a scheme in which mental states can be thought of as *outer* sentences, e.g.,

$$believe(Jon, publication(AAAI) = AIMag, because \ldots)$$

where ellipsis denotes the reasons for Jon's belief that *AI Magazine* is a publication of AAAI. The point of representing mental states with such sentences is that the grounds for having a belief can be included. The advantage gained by this is two-fold. In a belief revision system, when we are required to do belief revision, the incorporation of the reasons for having a belief simplifies our work. On the other hand, when we use beliefs as usual (i.e., no belief revision is required), we simply enter the related context and assert them. For example, in an outer context, the above sentence (about *AI Magazine*) with reasons is asserted. In an inner context, the simpler sentence $publication(AAAI) = AIMag$ would suffice, provided that we have committed ourselves to reasoning with this last proposition.

Guha on Contexts

Guha [30, 31] finds an essential use for formal contexts in implementing his so-called *micro-theories*. Micro-theories are theories of limited domains. Intuitively, micro-theories are the context's way of seeing the world, and are considered to have the following two basic properties: (i) there is a set of axioms related to each micro-theory, and (ii) there is a vocabulary which tells us the syntax and semantics of each predicate and each function specific to the micro-theory. Similar to McCarthy's conception, micro-theories are interrelated via lifting rules stated in an outer context.

Guha suggests several ways of using contexts effectively in reasoning, including the following:

- Contexts might be useful in putting together a set of related axioms. In this way, contexts are used as a means for referring to a group of related assertions (closed under entailment) about which something can be said.
- Contexts can be used as a mechanism for combining different theories. If the assertions in one context were not automatically available in other contexts, the system might as well be a set of disconnected knowledge bases. Therefore, by using lifting rules, different micro-theories may be integrated.

- Using contexts, we might have multiple models of a task. For example, regarding the task of finding out what to do in case of fire, we may offer different models for a workplace and for a house. In a workplace, the first thing to do may be to take away a file of documents, whereas, in a house, children must be saved first.

Lifting rules might be used to transfer facts from a (source) context to another (target) context. In the target context, the scope of quantifiers, the interpretation of objects, and even the vocabulary may change. Therefore, when we state a lifting rule, we must take all the possible outcomes into account. In the case of natural language, the problem becomes more complicated since indexicals and demonstratives come into play. Lifting rules should be definitely nonmonotonic. Guha uses default reasoning in the statement of lifting rules. His intuitions about the general lifting rules are as follows:

- *Default coreference*: Although there are differences among contexts, it can be expected that there will be similarities and overlaps. As a result, a significant number of terms in different contexts refer to (mean) the same thing. Such terms can be lifted from one context to another without any modification. Similarly, we can expect overlaps in many formulas, which may be lifted from one context to another without any change. Therefore, it will be a great simplification if we assume that a lifting operation will not require any modification, unless it is explicitly stated that there should be a change.

- *Compositional lifting*: Between contexts, there might be differences in vocabularies both in the words used and in the intended denotations of these words. In this case, specifying lifting rules for individual predicates should be enough for the system to use these rules in the lifting of formulas involving the predicates.

Although Guha's proposal accommodates any level of nesting on context, in CYC there are basically two levels: (i) micro-theories, and (ii) the default outer level. The lifting rules and general facts are stated in the outer level, and a problem is solved by the construction of a PSC under this level, unless the problem is local to a micro-theory.

S. Buvač et al. on Contexts

S. Buvač and Mason [17] (and in a more recent work, S. Buvač, V. Buvač, and Mason [16]) approach context from a mathematical viewpoint. They investigate the logical properties of contexts. They use the modality $\mathtt{ist}(c, p)$ to denote context-dependent truth and extend the classical propositional logic to what they call the *propositional logic of context.* (The quantificational logic of context is treated in [14].) In their proposal, each context is considered to have its own vocabulary—a set of propositional atoms which are defined (or meaningful) in that context.

S. Buvač and Mason discuss the syntax and semantics of a general propositional language of context, and give a Hilbert-style proof system for this language. The key contribution of their approach is providing a model theory for contexts. Two main results are the soundness and completeness proofs of this system. They also provide soundness and completeness results for various extensions of the general system, and prove that their logic is decidable.

The formal system is defined by the axioms (PL, K, and Δ) and inference rules (MP, Enter, Exit) given below:

(PL) $\vdash_c \phi$ (meaning: a formula ϕ is provable in context c (with a fixed vocabulary)) provided ϕ is an instance of a tautology.

(K) $\vdash_c \texttt{ist}(c_1, \phi \rightarrow \psi) \rightarrow (\texttt{ist}(c_1, \phi) \rightarrow \texttt{ist}(c_1, \psi))$ (meaning: every context is closed with respect to logical consequence).

(Δ) $\vdash_c \texttt{ist}(c_1, \texttt{ist}(c_2, \phi) \vee \psi) \rightarrow \texttt{ist}(c_1, \texttt{ist}(c_2, \phi)) \vee \texttt{ist}(c_1, \psi)$ (meaning: every context is "aware" of what is true in every other context).

(MP) From $\vdash_c \phi$ and $\vdash_c \phi \rightarrow \psi$ infer $\vdash_c \psi$.

(Enter) From $\vdash_{c'} \texttt{ist}(c, \phi)$ infer $\vdash_c \phi$; (Exit) From $\vdash_c \phi$ infer $\vdash_{c'} \texttt{ist}(c, \phi)$.

This system has the following two features [17]:

1. A context is modeled by a set of partial truth assignments which describe the possible states of affairs in that context. The `ist` modality is interpreted as validity: $\texttt{ist}(c, p)$ is true if and only if the propositional atom p is true in all the truth assignments associated with context c.

2. The nature of particular contexts is itself context dependent. The example of S. Buvač and Mason for this is Tweety, which has different interpretations when it is considered in a nonmonotonic reasoning literature context, and when it is considered in the context of *Tweety & Sylvester*. This observation leads us to consider a context as a sequence of individual contexts rather than a solitary context. In S. Buvač and Mason's terminology such a property is known as *non-flatness* of the system. The acceptance of a sequence of contexts respects the intuition that what holds in a particular context can depend on how this context is reached.[18]

S. Buvač and Mason show that the acceptance of the outermost context simplifies the meta-mathematics of the contexts. They first assume that there is no outermost context and build a proof system on this assumption. Then, they show that introducing the outermost context only simplifies the way they are dealing with non-flatness.

[18] The formal system presented earlier implicitly assumes flatness (every context looks the same regardless of which context it is being studied from) and would have be modified to deal with non-flatness. This is done by replacing the rules Enter and Exit with a single rule C [16].

F. Giunchiglia at al. on Contexts

F. Giunchiglia [23] takes a context to be a theory of the world which encodes an agent's perspective of it and which is used during a given reasoning process. A context is necessarily partial and approximate. Contexts are not situations (of situation calculus), for a situation is the complete state of the world at a given instant.

In formalizing context, the point of departure of F. Giunchiglia is partitioned databases. (Cf. [26] for origins of this work.) Each partition, A_i, may have different vocabulary. For example, while A_1 supports arithmetic operations, A_2 may support logical operations. With this approach, notion of well-formedness can be localized and may be distinct for each partition A_i. In formal terms, a context c_i is a triple $\langle L_i, A_i, \Delta_i \rangle$ where L_i is the language of the context, A_i is the axioms of the context, and Δ_i is the inference mechanism of the context. Under this definition, linking (bridge) rules are of the form

$$\frac{< A_i, c_i >}{< A_j, c_j >}$$

where A_i is a formula in c_i and A_j is the newly derived formula in c_j (also called a *justified assumption*). F. Giunchiglia offers the following to show the use of bridge rules:

- The usual Modus Ponens (MP) can be represented as

$$\frac{< A \rightarrow B, c_i > \;\; < A, c_i >}{< B, c_i >}$$

- A multi-contextual version of MP is represented as

$$\frac{< A \rightarrow B, c_i > \;\; < A, c_j >}{< B, c_k >}$$

- McCarthy's `ist` formula (asserted in c') becomes

$$\frac{< A, c >}{< \mathtt{ist}(c, A), c' >}$$

  (If we can prove A in context c, then we can prove in context c' that we can prove A in c.)

The first rule allows us to derive B inside a context just because we have derived $A \rightarrow B$ and A in the same context. Multi-contextual MP allows us to derive B in context c_k just because we have $A \rightarrow B$ derived in context c_i and A derived in context c_j. If these three contexts are assumed to represent the beliefs of three agents, then it is seen that B is not asserted as a result of deduction in c_k, but rather as a consequence of dissemination of results from c_i and c_j.

In a related work [25], F. Giunchiglia and Serafini formalize multilanguage (ML) systems of the above sort and propose them as an alternative to modal logic. (ML systems allow a hierarchy of first-order languages, each language containing names for the language below.) They then offer technical, epistemological, and implementation motivations to justify their proposal. Two useful applications of multilanguage systems can be found in [24, 22].

Attardi and Simi on Contexts

Attardi and Simi [7] offer a viewpoint representation which primarily depends on the view of context in a natural deduction sense. According to Attardi and Simi, contexts are sets of reified sentences of the FOL.

The main purpose of Attardi and Simi is to present a formalization of the notion of *viewpoint* as a construct meant for expressing varieties of relativized truth. The formalization is done in a logic which extends the FOL through an axiomatization of provability and with the proper reflection rules.[19]

The basic relation in the formalization is $\mathsf{in}('A', vp)$ where A is a sentence provable from viewpoint vp by means of natural deduction techniques. Viewpoints denote sets of sentences which represent the axioms of a theory. Viewpoints are defined as a set of reified meta-level sentences.

Since viewpoints are defined to be sets of reified sentences, operations between viewpoints are carried out via meta-level rules, e.g.,

$$\frac{\mathsf{in}('B', vp \cup \{'A'\})}{\mathsf{in}('A \rightarrow B', vp)}$$

This corresponds to the following in classical logic:

$$\frac{vp \cup \{A\} \vdash B}{vp \vdash A \rightarrow B}$$

The effective use of viewpoints in doing useful proofs requires a connection between the meta-level and the object-level rules. The rules below accomplish that:

$$\frac{vp_1 \vdash \mathsf{in}('A', vp_2)}{vp_1 \cup vp_2 \vdash A} \qquad \textit{(reflection)}$$

$$\frac{vp \vdash_C A}{\vdash \mathsf{in}('A', vp)} \qquad \textit{(reification)}$$

The notation $\vdash_C$ stands for "classically derivable" or "derivable without using the reflection rules."

Attardi and Simi cite a wide range of examples using the viewpoints. For instance, using viewpoints the notions of belief, knowledge, truth, and situation can be formalized as follows:

- *Belief*: The belief of an agent g is captured by means of in sentences, using $vp(g)$ as the viewpoint corresponding to the set of assumptions of the agent. Thus,

  $$\mathsf{Bel}(g, A) = \mathsf{in}(A, vp(g))$$

  and, by the reflection rule

  $$\mathsf{in}(A, vp(g)) \rightarrow (vp(g) \rightarrow A)$$

  we can use the beliefs of an agent.

[19] The reader is referred to [5] for an early sketch of the theory of viewpoints. This is basically an extension of FOL, adjoined by a carefully formulated reflective theory where mixed object/meta level statements are allowed.

- *Truth*: Truth is captured as provability in a special theory, viz. the Real World (RW). Ideally, everything that is true should be derivable in this theory, and truth can be defined as

$$\mathsf{True}(A) = \mathsf{in}(A, \mathsf{RW})$$

- *Knowledge*: Attardi and Simi view knowledge as true belief:

$$\mathsf{K}(g, A) \quad = \mathsf{Bel}(g, A) \wedge \mathsf{True}(A)$$

  Clearly, all the properties usually ascribed to knowledge can be derived, e.g., $\mathsf{K}(g, A) \rightarrow A$.

- *Situations à la Barwise/Perry*: Attardi and Simi take situations as sets of basic facts [11], and use an approach similar to that of belief. Thus, they define a basic relation

$$\mathsf{Holds}(A, s) = \mathsf{in}(A, vp(s))$$

  where $vp(s)$ is the set of facts holding in a situation s. (See the next section for more information on situations.)

Shoham on Contexts

Shoham [56] uses the alternative notation p^c to denote that assertion p holds in context c. According to him, every assertion is meaningful in every context, but the same assertion might have different truth values in different contexts.[20] Thus, his approach is quite different compared to the approaches of McCarthy, Guha, and S. Buvač et al.

Shoham describes a propositional language depending on his *more general than* relation ($\dot{\supset}$). The relation defines a weak partial ordering between contexts; not every pair of contexts is comparable under it. Is there a most general (or most specific) context? Mathematically this corresponds to the question "Is there an upper (or lower) bound on $\dot{\supset}$?" In Shoham's proposal, the question is not answered, but when the system is analyzed the existence of the most general and the most specific contexts is considered.[21]

The language Shoham describes is quite similar to FOL but his relations $\dot{\supset}, \dot{\vee}, \dot{\wedge}$, and $\dot{\neg}$ work over contexts. Here, $x \dot{\wedge} y$ is defined as the greatest lower bound on x and y with respect to $\dot{\supset}$ (if it exists). Similarly, $x \dot{\vee} y$ is defined as a least upper bound of the contexts x and y (if it exists). When defined, $\dot{\neg} x$ is the context which

[20] Shoham notes that "a general treatment of contexts may indeed wish to exempt contexts from the obligation to interpret every assertion" [56, p. 400].

[21] Shoham gives a particularly nice account of the term "more general than" [56, p. 398]: "[W]hether one of the two contexts *human beings in general* and *human beings in conditions where influenza viruses are present* is more general than the other depends on what we mean by each one. If the first context includes some information about people and the second context includes that information plus further information about viruses, then the former is more general. If the first includes *all* information about people and the second some subset of it that has to do with viruses, then the latter is more general. Otherwise, the two are noncomparable."

is not comparable[22] to x under $\dot{\supset}$. A context set is *and-closed* if it is closed under conjunction, *or-closed* if it is closed under disjunction, *and-or-closed* if it is both, *not-closed* if it is closed under negation, and simply *closed* if it is all three. From these definitions, we see that if an or-closed context set contains both x and $\dot{\neg}x$ for some x, then the context set contains the most general context, i.e., the *tautological* context. Similarly, under the same condition, an and-closed context set contains the most specific context, i.e., the *contradictory* context.

What should be the logical status of p^c? Shoham takes it to be an assertion and introduces a simple language for discussing contexts and assertions about them. Basically, given a context set C with the partial order $\dot{\supset}$ and a propositional language $\mathcal{L}$, the set of well-formed formulas is the smallest set S such that

1. If $c_1, c_2 \in C$ then $c_1 \dot{\supset} c_2 \in S$.
2. If $p \in \mathcal{L}$ then $p \in S$.
3. If $s \in S$ and $c \in C$ then $s^c \in S$.
4. If $s_1, s_2 \in S$ then $s_1 \wedge s_2 \in S$ and $\neg s_1 \in S$.

Shoham's purpose is not really to offer "The Right Semantics" for p^c; he is more interested in identifying some options towards this ultimate goal and investigating the interaction between modal operators (e.g., the knowledge operator K in the logic of knowledge and belief) and context. Some interesting proposals are given in this direction to investigate the notion of *contextual knowledge*, i.e., the meaning of $\mathtt{K}^c p$—his notation for "p is known in context c."

The Situation Theoretic Approach

Background

The standard reference on Situation Theory is Devlin's excellent book [20]. The original work of Barwise and Perry [12] is still worthy of study and remains as an elegant philosophical argument for introducing situations. Unfortunately, the notation and (sometimes) the terminology of the latter reference are rather outdated. Accordingly, we use Devlin's notation and terminology in the following.

According to Situation Theory, infons are the basic informational units (discrete items of information). They are denoted as $\ll P, a_1, \ldots, a_n, i \gg$ where P is an n-place relation, $a_1, \ldots, a_n$ are objects appropriate for the respective argument places of P, and i is the polarity (1 or 0) indicating whether the relation does or does not hold.

Situations are first-class citizens of the theory, and are defined intensionally. A situation is considered to be a structured part of the reality that an agent manages to pick out and/or individuate. Situations and infons are related by the *supports* relation:

[22] The term *not comparable* means that each context contains some axioms which are not contained in the other context; cf. the preceding note.

s supports α (denoted $s \models \alpha$) means that α is an infon that is true of s.

For example, a situation s in which Bob hugs Carol would be described by $s \models \ll hugs, Bob, Carol, l, t, 1 \gg$ where l and t together give the spatio-temporal coordinates of this hugging action.

Abstract situations are the constructs which are more amenable to mathematical manipulation. An abstract situation is defined as a (possibly non-well-founded [11]) set of infons. Given a real situation s, the set $\{\alpha \mid s \models \alpha\}$ is the corresponding abstract situation.

One of the important ideas behind Situation Theory is a *scheme of individuation*, a way of carving the world into uniformities. Being constructs that link the scheme of individuation to the technical framework of the theory, types are important features of Situation Theory. Just as individuals, temporal locations, spatial locations, relations, and situations, types are also uniformities that are discriminated by agents. Relations may have their argument places filled either with individuals, situations, locations, and other relations, or with types of individuals, situations, locations, and relations. Some *basic types* are TIM (the type of a temporal location), LOC (the type of a spatial location), IND (the type of an individual), SIT (the type of a situation), etc.

In Situation Theory, for each type T, an infinite collection of *basic parameters* $T_1, T_2, \ldots$ is introduced. For example IND_3 is an IND-parameter. We use the notation $\dot{l}$, $\dot{t}$, $\dot{a}$, $\dot{s}$, etc. to denote parameters of type LOC, TIM, IND, SIT, etc., respectively. Sometimes, rather than parameters ranging over all individuals, we need parameters that range over a more restricted class, viz. *restricted parameters*. For example,

$$\dot{r}1 = \dot{a} \upharpoonright \ll kicking, \dot{a}, \dot{b}, 1 \gg$$
$$\dot{a} = IND3 \upharpoonright \ll man, IND3, 1 \gg$$
$$\dot{b} = IND2 \upharpoonright \ll football, IND2, 1 \gg$$

In this case, $\dot{r}1$ ranges over all men kicking footballs.

We use the term *parametric infon* to emphasize that in a particular infon one or more parameters occur free. Infons that have no free parameters are called *parameter-free*. Related to parametric infons, there is a construct by which we can assign values to parameters. We call this an *anchor*. Formally, an anchor for a set, A, of basic parameters is a function defined on A, which assigns to each parameter T_i in A an object of type T. Therefore, if f is an anchor for A and T_i is a parameter in A, then

$$\ll \textit{of-type}, f(T_i), T, 1 \gg$$

For example, if f anchors $\dot{a}$ to the type IND_3 individual "Sullivan," we write

$$f(\dot{a}) = \text{Sullivan}$$

to denote this anchoring.

Let s be a given situation. If $\dot{x}$ is a parameter and I is a set of infons (involving $\dot{x}$), then there is a type

$$[\dot{x}|s \models I]$$

This is the type of all those objects to which $\dot{x}$ may be anchored in s, for which the conditions imposed by I obtain. We refer to this process of obtaining a type from a parameter $\dot{x}$, a situation s, and a set I of infons, as *type-abstraction*. $\dot{x}$ is known as the *abstraction parameter* and s is known as the *grounding situation*.

In Situation Theory, the flow of information is realized via *constraints*, represented as

$$S_0 \Rightarrow S_1$$

Here S_0 and S_1 are situation-types. Cognitively, if this relation holds, then it is a fact that if S_0 is realized (i.e., there is a real situation $s_0 : S_0$), then so is S_1 (i.e., there is a real situation $s_1 : S_1$). For example, with the constraint $S_s \Rightarrow S_f$, we might represent the regularity "Smoke means fire", provided that we have

$$S_s = [\dot{s}|\dot{s} \models \ll \textit{smoke-present}, \dot{l}, \dot{t}, 1 \gg]$$
$$S_f = [\dot{s}|\dot{s} \models \ll \textit{fire-present}, \dot{l}, \dot{t}, 1 \gg]$$

This constraint is read as "S_s involves S_f," and represents a fact (i.e., a factual, parameter-free infon):

$$\ll involves, S_s, S_f, 1 \gg$$

Attunement to this constraint is what enables an intelligent agent that sees smoke in a situation to realize that there is a fire.

Barwise on Contexts

Barwise's ideas on circumstance, thus on context, are best articulated in his work on conditionals and circumstantial information [8] and this is the essential source we will review here. The attentive reader has presumably noticed that situations represent a way of modeling contexts. In fact, in [10] Barwise expounds why a context is a situation. Briefly, he proposes that there is a definite relationship between situations and what are known as *context sequences* in possible world semantics [2]. In possible world semantics, given a sentence s, the person p who uttered the sentence, the spatio-temporal location of the utterance l and t, the object o that p is referring to, etc. are all lumped together into a sequence $c = \langle p, l, t, o, \ldots \rangle$. Basically, c represents various contextual elements that play a role in obtaining the propositional content of any particular use of s. Barwise claims that c is nothing more than a representation of a situation, that portion of the world that is essentially needed (relevant) to the determination of the content of the utterance of s. Thus, his claim is that by admitting situations, one no longer needs ad hoc devices like c.

The Missing Pollen

Let us consider Claire (Barwise's then nine-month old daughter). Barwise knows that if Claire rubs her eyes, then she is sleepy. This is expressed by the conditional statement

If Claire rubs her eyes, then she is sleepy.

For months, this was a sound piece of (conditional) knowledge that Barwise and his wife used to understand Claire, and learn when they should put her to bed. However, in the early summer, it began to fail them. Combined with other symptoms, Barwise and his wife eventually figured out that Claire was allergic to something or other. They called it Pollen X since they did not know its precise identity. So Pollen X could also cause Claire to rub her eyes.

Barwise formalizes the problem stated in the above example as follows. Briefly, with constraint $C = [S \Rightarrow S']$, a real situation s contains information relative to such an actual constraint C, if $s : S$. Clearly, s may contain various pieces of information relative to C, but the most general proposition that s contains, relative to C, is that s' is realized, where $s' : S'$.

Thus we can represent the conditional information above with the following parametric constraint C:

$$\begin{aligned} S &= [\dot{s} | \dot{s} \models \ll rubs, Claire, eyes, \dot{l}, \dot{t}, 1 \gg] \\ S' &= [\dot{s} | \dot{s} \models \ll sleepy, Claire, \dot{l}, \dot{t}, 1 \gg] \\ C &= [S \Rightarrow S'] \end{aligned}$$

Before Pollen X was present, the above constraint represented a reasonable account. However, when Pollen X arrived, the constraint became inadequate and required revision. Barwise points out to two alternatives to deal with the problem:

- From [if ϕ then ψ] infer [if ϕ and β, then ψ].
- From [if ϕ then ψ] infer [if β, then if ϕ then ψ].

Here β corresponds to the additional background conditions.

Barwise chooses the second way,[23] modifies *involves*, and makes the background assumptions explicit by introducing a third parameter B:

$$S_0 \Rightarrow S_1 | B$$

With the new *involves*, the Missing Pollen Example can be solved via the introduction of a B which supports the following:

$$\ll exists, PollenX, \dot{l}, \dot{t}, 0 \gg$$

[23]Although, these alternatives are equivalent from a logical point of view, the second is more appropriate to reflect the intuitions behind the background conditions. In the first case, the rule *if ϕ then ψ* is directly modified to use background conditions, whereas, in the second case, it is not touched, but is evaluated only when the background conditions hold. Introduction of background conditions for rules corresponds to a nonmonotonic reasoning mechanism.

$$S = [\dot{s}|\dot{s} \models \ll rubs, Claire, eyes, \dot{l}, \dot{t}, 1 \gg]$$
$$S' = [\dot{s}|\dot{s} \models \ll sleepy, Claire, \dot{l}, \dot{t}, 1 \gg]$$
$$B = [\dot{s}|\dot{s} \models \ll exists, PollenX, \dot{l}, \dot{t}, 0 \gg]$$
$$C = [S \Rightarrow S'|B]$$

In the beginning it was winter and there were no pollen. The context, call it c_1, must be a situation type which supports

$$c_1 \models \ll exists, PollenX, \dot{l}, \dot{t}, 0 \gg$$

(and possibly other things related to Claire, rubbing one's eyes, etc.). Using context c_1 as the grounding situation, we do not violate the background condition B of constraint C , and thus can conclude that "Claire is sleepy."

Later, in summer, the new context, c_2, supports the infon

$$c_2 \models \ll exists, PollenX, \dot{l}, \dot{t}, 1 \gg$$

and when we use c_2 as the grounding situation, we are faced with an inconsistency between B and c_2. Therefore, C becomes void in the new context of the talk, and the conclusion "Claire is sleepy" cannot be reached.

I am a philosopher

We will prove that the content of all three sentences are the same, i.e., A is a philosopher.

We have three contexts associated with each individual in the conversation: c_A, c_B, and c_C, respectively. We will represent the indexicals with special parameters $\dot{I}, \dot{You}$, and $\dot{She}$.

In c_A, we have the following infons supported:

$$\ll corresponds, \dot{I}, A, 1 \gg$$
$$\ll philosopher, \dot{I}, 1 \gg$$

where *corresponds* is a function which associates an indexical to a person, and utterances about being a philosopher are represented with infons of type $\ll philosopher, \dot{x}, 1 \gg$.

c_B supports the following:

$$\ll corresponds, \dot{She}, A, 1 \gg$$
$$\ll philosopher, \dot{She}, 1 \gg$$

c_C supports:

$$\ll corresponds, \dot{You}, A, 1 \gg$$
$$\ll philosopher, \dot{You}, 1 \gg$$

Now, it is a trivial matter to observe that $\dot{I}, \dot{You}$, and $\dot{She}$ all collapse to A, for the anchoring

$$f(\dot{I}) = A$$
$$f(\dot{You}) = A$$
$$f(\dot{She}) = A$$

does the job. Consequently, the utterance of A might be decontextualized as $\ll philosopher, A, 1 \gg$.

The Obligatory Tweety Example

As we have stated before, in Situation Theory, we represent implications with constraints. While stating the constraints, we can use background conditions to add nonmonotonicity:

$$S_0 = [\dot{s}|\dot{s} \models \ll bird, \dot{x}, 1 \gg]$$
$$S_1 = [\dot{s}|\dot{s} \models \ll flies, \dot{x}, 1 \gg]$$
$$B = [\dot{s}|\dot{s} \models \ll penguin, \dot{x}, 0 \gg \wedge \dot{s} \models \ll present, Air, 1 \gg]$$
$$C = [S_0 \Rightarrow S_1 | B]$$

The constraint C states that every bird flies unless it is a penguin or there is no air. Here, the important contribution of the situation-theoretic account is that the environmental factors can be easily included in the reasoning phase by suitably varying B.

Conclusion

The comparison of the previous approaches is summarized in Table 1 where the first row marks the language of formalization. In order to have a less populated table, not all the pertinent works (by an author or a group of authors) have been listed and instead an exemplar publication has been selected as a starting point. It is noted that except for Barwise, all of the previous approaches are stated in a more or less logicist framework. Among these, Shoham, and S. Buvač and Mason propose context as a modal operator; the latter further considers contextual reasoning in a natural deduction sense, and allows operations of entering/exiting contexts. F. Giunchiglia and coworkers propose their multilanguage systems as a true alternative to modal logic.

Among the previous approaches, McCarthy's and Guha's are not paradox free, whereas S. Buvač and Mason's, and Attardi and Simi's approaches are paradox free. Shoham's approach must be paradox free: it is basically propositional modal logic. We do not know whether Barwise's (hence our) approach is paradox free or not. However, in a thought-provoking work [11], Situation Theory is shown to be powerful enough to deal with circularity. The last row of the table reflects this. (It

	Mc93	Gu91	Sh91	Gi93	BM93	AS95	Ba86
Logic vs. Situation Theory	Logic	Logic	Logic	Logic	Logic	Logic	S.T.
Modal Treatment	No	No	Yes	No	Yes	No	No
Natural Deduction	Yes	Yes	No	Yes	Yes	Yes	No
Paradox Free	No	No	Yes	Yes	Yes	Yes	?
Circularity	No	No	No	No	No	Yes	Yes

Legend:

Mc93	Notes on formalizing context [45]
Gu91	Contexts: A formalization and some applications [30]
Sh91	Varieties of context [56]
Gi93	Contextual reasoning [23]
BM93	Propositional logic of context [17]
AS95	A formalization of viewpoints [7]
Ba86	Conditionals and conditional information [8]

Table 1: Comparison of approaches to formalizing context.

must be noted that an important portion of Attardi and Simi's work is also focused on evading paradoxes.)

Clearly, other tables can be made to highlight the motivations, expressiveness, and complexity[24] of particular approaches. Rather than doing this, we simply state that in the end, applicability will be the sole basis of ratification for any approach. While the idea of formalizing context seems to have caught on and produced good theoretical outcomes, the area of innovative applications remains relatively unexplored. This is evidently the place to which further research on context should converge.

Acknowledgments

A preliminary version of this paper was submitted to this Magazine in October, 1994. It then underwent a revision and we are grateful to the Editor and an anonymous referee for thoughtful remarks and suggestions. For the penultimate version of the paper we have benefited from the very detailed and constructive criticism of two anonymous referees and the Editor Emeritus, and once again, we are thankful. As usual, responsibility for the contents remains with us.

[24] We refer the reader to Massacci [41] for some complexity results regarding contextual reasoning.

References

[1] V. Akman and M. Surav. Contexts, oracles, and relevance. In S. Buvač, editor, *AAAI Fall Symposium on Formalizing Context*, AAAI Technical Report Series, No. FS-95-02, pages 23–30. AAAI Press, Menlo Park, CA, 1995.

[2] V. Akman and E. Tın. What is in a context? In E. Arıkan, editor, *Proceedings of the 1990 Bilkent International Conference on New Trends in Communication, Control, and Signal Processing*, pages 1670–1676. Elsevier, Amsterdam, 1990.

[3] J. Allen. *Natural Language Understanding*. The Benjamin/Cummings Publishing Company, Inc., Redwood City, CA, 1995. Second Edition.

[4] P. A. Angeles. *Dictionary of Philosophy*. Harper and Row Publishers, New York, 1981.

[5] G. Attardi and M. Simi. Reflections about reflection. In J. A. Allen, R. Fikes, and E. Sandewall, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Second Conference*, Morgan Kaufmann, San Mateo, CA, 1991.

[6] G. Attardi and M. Simi. Proofs in context. In J. Doyle and P. Torasso, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourth International Conference*, Morgan Kaufmann, San Mateo, CA, 1994.

[7] G. Attardi and M. Simi. A formalization of viewpoints. *Fundamenta Informaticae*, 23(3):149–174, 1995. Also technical report TR-93-062, International Computer Science Institute, University of California, Berkeley, CA, 1993.

[8] J. Barwise. Conditionals and conditional information. In E. C. Traugott, C. A. Ferguson, and J. S. Reilly, editors, *On Conditionals*, pages 21–54. Cambridge University Press, Cambridge, UK, 1986.

[9] J. Barwise. On the circumstantial relation between meaning and content. In *The Situation in Logic*, number 17 in CSLI Lecture Notes, pages 59–77. Center for the Study of Language and Information, Stanford, CA, 1987.

[10] J. Barwise. Situations and small worlds. In *The Situation in Logic*, number 17 in CSLI Lecture Notes, pages 79–92. Center for the Study of Language and Information, Stanford, CA, 1987.

[11] J. Barwise and J. Etchemendy. *The Liar: An Essay on Truth and Circularity*. Oxford University Press, New York, 1987.

[12] J. Barwise and J. Perry. *Situations and Attitudes*. MIT Press, Cambridge, MA, 1983.

[13] J. Barwise and J. Seligman. The rights and wrongs of natural regularity. In J. Tomberlin, editor, *Philosophical Perspectives 8: Logic and Language*. Ridgeview Press, Atascadero, CA, 1992.

[14] S. Buvač. Quantificational logic of context. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, 1996.

[15] S. Buvač. Resolving lexical ambiguity using a formal theory of context. In K. van Deemter and S. Peters, editors, *Semantic Ambiguity and Underspecification*, number 55 in CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, CA, 1996.

[16] S. Buvač, V. Buvač, and I. A. Mason. Metamathematics of contexts. *Fundamenta Informaticae*, 23(3):263–301, 1995.

[17] S. Buvač and I. A. Mason. Propositional logic of context. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pages 412–419, Washington, DC, 1993.

[18] E. Charniak. Motivation analysis, abductive unification, and nonmonotonic equality. *Artificial Intelligence*, 34:275–296, 1988.

[19] H. H. Clark and T. B. Carlson. Context for comprehension. In J. Long and A. Baddeley, editors, *Attention and Performance IX*, pages 313–330. Lawrence Erlbaum Associates, Hillsdale, NJ, 1981.

[20] K. Devlin. *Logic and Information.* Cambridge University Press, New York, 1991.

[21] T. J. Froehlich. Relevance reconsidered—towards an agenda for the 21st century: Introduction to special topic issue on relevance research. *Journal of the American Society for Information Science*, 45:124–134, 1994.

[22] E. Giunchiglia, P. Traverso, and F. Giunchiglia. Multi-context systems as a specification framework for complex reasoning systems. In *Proceedings of the ECAI-92 Workshop on Formal Specification Methods for Complex Reasoning Systems*, Vienna, 1992.

[23] F. Giunchiglia. Contextual reasoning. *Epistemologia*, XVI:345–364, 1993. Special Issue on *I Linguaggi e le Macchine*; Short version in *Proceedings of the IJCAI'93 Workshop on Using Knowledge in its Context*, Chambery, France, 1993.

[24] F. Giunchiglia and L. Serafini. Multilanguage first order theories of propositional attitudes. In *Proceedings of the Third Scandinavian Conference on Artificial Intelligence*, pages 228–240, IOS Press, Amsterdam, 1991.

[25] F. Giunchiglia and L. Serafini. Multilanguage hierarchical logics, or: How we can do without modal logics. *Artificial Intelligence*, 65(1):29–70, 1994.

[26] F. Giunchiglia and R. W. Weyhrauch. A multi-context monotonic axiomatization of inessential nonmonotonicity. In P. Maes and D. Nardi, editors, *Meta-Level Architectures and Reflection*, pages 271–285. Elsevier, Amsterdam, 1988.

[27] C. H. Goh, S. T. Madnick, and M. Siegel. Ontologies, contexts, and mediation: Representing and reasoning about semantic conflicts in heterogeneous and autonomous systems. Working Paper 3848, Sloan School of Management, MIT, Cambridge, MA, 1995.

[28] C. Goodwin and A. Duranti. Rethinking context: An introduction. In A. Duranti and C. Goodwin, editors, *Rethinking Context: Language as an Interactive Phenomenon*, pages 1–42. Cambridge University Press, New York, 1992.

[29] B. J. Grosz and C. L. Sidner. Attention, intention, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.

[30] R. V. Guha. *Contexts: A Formalization and Some Applications.* PhD thesis, Computer Science Department, Stanford University, Stanford, CA, 1991. Also technical report ACT-CYC-423-91, Microelectronics and Computer Technology Corporation, Austin, TX, 1991.

[31] R. V. Guha. Context dependence of representations in CYC. Technical Report CYC 066-93, Microelectronics and Computer Technology Corporation, Austin, TX, 1993.

[32] R. V. Guha and D. B. Lenat. CYC: A midterm report. *AI Magazine*, 11(3):32–59, 1990.

[33] S. P. Harter. Psychological relevance and information science. *Journal of the American Society for Information Science*, 43:602–615, 1992.

[34] M. A. Hearst. *Context and Structure in Automated Full-Text Information Access.* PhD thesis, Computer Science Division, University of California, Berkeley, CA, 1994.

[35] J. R. Hobbs, M. Stickel, D. Appelt, and P. Martin. Interpretation as abduction. *Artificial Intelligence*, 63(1–2):69–142, 1993.

[36] H. Kamp and U. Reyle. *From Discourse to Logic: Introduction to Model-Theoretic Semantics of Natural Language, Formal Logic, and Discourse Representation Theory.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.

[37] D. Kaplan. Demonstratives. In J. Almog, J. Perry, and H. Wettstein, editors, *Themes from Kaplan*, pages 481–563. Oxford University Press, New York, 1989.

[38] G. Leech. *Semantics: The Study of Meaning.* Penguin Books, 1981. Second Edition.

[39] D. B. Lenat. CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11):33–38, 1995.

[40] J. Lyons. *Linguistic Semantics: An Introduction.* Cambridge University Press, Cambridge, UK, 1995.

[41] F. Massacci. Contextual reasoning is NP-complete. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, 1996.

[42] J. McCarthy. Some expert systems need common sense. In H. Pagels, editor, *Computer Culture: The Scientific, Intellectual, and Social Impact of the Computer*, pages 129–137. Annals of the New York Academy of Sciences, Vol. 426, 1984.

[43] J. McCarthy. Generality in artificial intelligence. *Communications of the ACM*, 30(12):1030–1035, 1987.

[44] J. McCarthy. Artificial intelligence, logic and formalizing common sense. In R. H. Thomason, editor, *Philosophical Logic and Artificial Intelligence*, pages 161–190. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1989.

[45] J. McCarthy. Notes on formalizing context. In *Proceedings of the Thirteenth International Joint Conference in Artificial Intelligence*, pages 555–560, Chambery, France, 1993.

[46] J. McCarthy. A logical AI approach to context. Technical note, Computer Science Department, Stanford University, Stanford, CA, 1995.

[47] J. McCarthy and S. Buvač. Formalizing context (expanded notes). Technical Note STAN-CS-TN-94-13, Computer Science Department, Stanford University, Stanford, CA, 1994.

[48] P. P. Nayak. Representing multiple theories. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, 1994.

[49] T. K. Park. Toward a theory of user-based relevance: A call for a new paradigm of inquiry. *Journal of the American Society for Information Science*, 45:135–141, 1994.

[50] W. V. O. Quine. Propositional objects. In *Ontological Relativity and Other Essays*. Columbia University Press, New York, 1969.

[51] F. Recanati. *Direct Reference: From Language to Thought*. Blackwell Publishers, Oxford, UK, 1993.

[52] M. P. Reddy and A. Gupta. Context interchange: A lattice based approach. *Knowledge-Based Systems*, 8(1):5–13, 1995.

[53] E. Rosch. Principles of categorization. In E. Rosch and B. B. Lloyd, editors, *Cognition and Categorization*, chapter 2. Lawrence Erlbaum Associates, Hillsdale, NJ, 1978.

[54] R. C. Schank and C. J. Rieger. Inference and computer understanding of natural language. *Artificial Intelligence*, 5:373–412, 1974.

[55] J. Seligman. Structure of regularity. Draft Lecture Notes, Indiana University Logic Group, Bloomington, IN, 1993.

[56] Y. Shoham. Varieties of context. In V. Lifschitz, editor, *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy*, pages 393–408. Academic Press, Boston, MA, 1991.

[57] E. Shortliffe. *MYCIN: Computer-Based Medical Consultations.* Elsevier, New York, 1976.

[58] D. Sperber and D. Wilson. *Relevance: Communication and Cognition.* Basil Blackwell, Oxford, UK, 1986.

[59] M. Surav and V. Akman. Modeling context with situations. In P. Brezillon and S. Abu-Hakima, editors, *Working Notes of the IJCAI-95 Workshop on Modeling Context in Knowledge Representation and Reasoning*, pages 145–156. 1995. Also technical report LAFORIA 95/11, Laboratoire Formes et Intelligence Artificielle, Institut Blaise Pascal, Université Paris VI, Paris, 1995.

[60] E. Tın and V. Akman. Computing with causal theories. *International Journal of Pattern Recognition and Artificial Intelligence*, 6(4):699–730, 1992.

[61] J. van Eijck and H. Kamp. Representing discourse in context. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Linguistics.* Elsevier, Amsterdam, 1996. To appear.

[62] L. Wittgenstein. *Philosophical Investigations.* Basil Blackwell, Oxford, UK, 1984. First published in 1953; translated by G. E. M. Anscombe.