Computational Intelligence, Volume 12, Number 4, 1996

THE USE OF SITUATION THEORY IN CONTEXT MODELING

VAROL AKMAN

Department of Computer Engineering and Information Science, Bilkent University, Bilkent, Ankara 06533, Turkey

MEHMET SURAV

NewNet, Inc., 2 Enterprise Drive, Shelton, Connecticut 06484, USA

At the heart of natural language processing is the understanding of context dependent meanings. This paper presents a preliminary model of formal contexts based on situation theory. It also gives a worked-out example to show the use of contexts in lifting, i.e., how propositions holding in a particular context transform when they are moved to another context. This is useful in NLP applications where preserving meaning is a desideratum.

Key words: context, lifting, knowledge representation, commonsense reasoning, natural language processing, situation theory.

1. INTRODUCTION

Our long-term goal is to offer a full-fledged formalization of context, one that can be used, among other areas, in natural language processing. In this paper, which can be regarded as a tentative step towards that end, we will identify the role of context in language and take a look at some salient efforts in (logical) AI treating formal contexts. In general, the focus of our discussion will be McCarthy's proposal (McCarthy 1987, 1993), which is the groundwork for all ensuing logicist formalizations. While the main purpose of McCarthy, viz. a mechanism by which we can build AI systems which are not forever stuck with the concepts they use at a given time (because they can surpass the context they are in), is still largely unfulfilled, we will see that important advances have been made (Guha 1991; Shoham 1991; McCarthy and S. Buvač 1994; S. Buvač and Mason 1995; Attardi and Simi 1995).

Our model of context, on the other hand, is inspired by the pioneering work of Barwise (1986) on conditionals, and will be presented using the notation and terminology of situation theory (Barwise and Perry 1983; Devlin 1991). After giving the minimum background to situation theory, we will state this model and discuss an application of it in lifting, i.e., the process of computing what is true in one context based on what is true in another context (McCarthy and S. Buvač 1994). The situation theoretic model has notable properties such as partiality, dynamic contexts, and natural language support. It links our work with NLP in a simple, natural way: reasoning is essential for NLP, knowledge representation is a prerequisite for any kind of reasoning, and situation theory can be used both as a KR scheme and to support contextual reasoning.[1]

[1]The reader is referred to (Akman and Surav 1996) for a better appreciation of formal contexts. There is little overlap between that survey and the present paper. A preliminary version of this paper has appeared as (Surav and Akman 1995). Also see (Akman and Surav 1995) for related ideas on contexts and situation theory.

2. CONTEXT IN NATURAL LANGUAGE

"You shall know a word by the company it keeps." This remark of J. R. Firth, famed British linguist, seems to us an apt reminder for the ubiquity of context. According to Crystal (1991, p. 78), 'context' is a general term in linguistics and phonetics to refer to specific parts of an utterance (or text) near or adjacent to a unit (e.g., a sound, word) which is the focus of attention. The occurrence of a unit is partly or wholly determined by its context, which is specified in terms of the unit's relations. Blackburn (1994, p. 80) offers a similar definition: "In linguistics, context is the parts of an utterance surrounding a unit and which may affect both its meaning and its grammatical contribution." However, he is quick to add that context also refers to "the wider situation, either of the speaker or of the surroundings, that may play a part in determining the significance of a saying."

Leech (1981, pp. 66–67) notes that the specification of context (whether linguistic or nonlinguistic) has the effect of narrowing down the communicative possibilities of a message. This particularization of meaning can take place in assorted ways, including: (i) elimination of certain ambiguities or multiple meanings in the message, (ii) clarification of the referents of deictics and definite descriptions, (iii) supplying of information which the writer has omitted through ellipsis, (iv) interpretation of tense, and (v) determination of the scope of quantifiers.

It is standard nowadays to use the term 'co-text' for the narrow, purely linguistic context (Lyons 1995, p. 271). As for the total nonlinguistic background to an utterance (including: the immediate situation in which it is used, the knowledge of speaker and hearer about the commonsense world, the knowledge of what has been said earlier, the relevant beliefs and presuppositions of speaker and hearer), the term 'situational context' has been offered (Crystal 1991, p. 79). Similarly, Lyons (1995, p. 271) uses the term 'context of situation' as a synonym for situational context. He believes that natural language meaning must be studied as a multiple phenomenon, its numerous aspects being relatable to (i) different levels of linguistic analysis, and (ii) features of the world.

Being one of those linguistic abstractions which is constantly used in all kinds of contexts [sic] but never explained, the establishment of relevant context in NLP is traditionally seen as a formidable problem. M. Pinkal voices this difficulty in a vivid passage that appears in (Asher and Simpson 1994, p. 733):

> Aside from the surrounding deictic coordinates, aside from the immediate linguistic co-text and accompanying gestural expressions at closer view, the following determinants can influence the attribution of sense: the entire frame of interaction, the individual biographies of the participants, the physical environment, the social embedding, the cultural and historical background, and—in addition to all these—facts and dates no matter how far removed in dimensions of time and space. Roughly speaking, 'context' can be the whole world in relation to an utterance act.

But observations such as these also show that the quest for idealized, context-independent meaning (which goes under the name 'logical form' in semantics) is seriously misdirected.[2] Taking an engineering attitude and studying contexts as

[2] Barwise (1986, p. 99) dubs the idea behind this quest the 'fleshing out strategy,' because it is based on the following assumption: sentences whose interpretation depends on some annoying contextual element can be fleshed out to sentences where that contextual element is eradicated. He then adds: "I assume that this strategy is wrong-headed, that it has been shown to be unworkable, and that it should now be laid to rest."

mathematical entities with properties useful in AI or NLP (to take two examples) seem to be the only viable approach. This is what McCarthy did in his distinguished work on formal contexts and what others (including us) hope to follow suit.

3. CONTEXT IN LOGICAL AI

3.1. The work of McCarthy

McCarthy offers no definition of context. His underlying assumption is that "[t]here are mathematical context structures of different properties, some of which are useful" (McCarthy 1996, p. 2). He wittily remarks that asking what a context is is like asking what a group element is.

McCarthy's basic relation relating contexts and propositions is $ist(c, p)$, asserting that proposition p is true in context c. The origin of this relation can be traced to the Turing lecture (the written version) of McCarthy (1987). McCarthy notes that formulas $ist(c, p)$ are always considered as themselves asserted within an outer context c_0 such that $ist(c_0, ist(c, p))$.

The importance of $ist(c, p)$ for NLP can be seen by noticing that while the set of propositions true in a context may be finite, the collection of natural language sentences that can express these propositions will be infinite (McCarthy 1996). This is especially crucial in translation tasks, e.g., interpreting a source language sentence and constructing an equivalent target language sentence. For typical translation tasks, the target sentence must succeed in communicating the propositional content of the original sentence, and having propositions (rather than sentences) as the basic building blocks helps in this endeavor (Farwell and Helmreich 1995).

Contexts are 'first-class citizens.' We can use contexts in our logical formulas in the same way we use other objects. In other words, contexts are formal objects in the semantics: they can be denoted by constants in the logical language and when necessary, variables can range over them.

Each context has a vocabulary associated with it. Thus, a given statement may not be expressible in some context (due to the impoverished vocabulary of that context). In yet another context it would be expressed differently. It is noted that since there is no absolute outermost context, it is necessary to have an adequate notion of transcendence, i.e., outstripping the outermost context so far referred to. Transcendence is the way to relax or modify some assumptions of an old context; it is essentially a move from a context that makes certain assumptions to one that does not (McCarthy and S. Buvač 1994).

In order to implement transcendence, an appropriate set of nonmonotonic rules for lifting sentences to broader contexts is required. By lifting a predicate (or formula, axiom, etc.) from one context to another related context, we mean transferring that predicate (or formula, axiom, etc.) to broader contexts—those involving fewer assumptions. As an illustration of lifting, consider the relation 'more general than' ($\preceq$). The inequality $c_1 \preceq c_2$ states that c_2 is more general than c_1 (equivalently, c_1 is a specialization of c_2). Essentially, c_2 involves no more assumptions than c_1. Using $\preceq$, a fact from a context to one of its super-contexts can be lifted via the lifting rule

$$\forall c_1 \forall c_2 \forall p \ (c_1 \preceq c_2) \wedge ist(c_1, p) \wedge \neg ab(p, c_1, c_2) \rightarrow ist(c_2, p)$$

where p is a proposition of c_1 and ab is an abnormality predicate to support nonmonotonicity.

When we regard contexts in the natural deduction sense—as McCarthy (1987) suggested—the operations of entering and leaving a context might be given succinct definitions (McCarthy and S. Buvač 1994). Basically, since $ist(c,p)$ will be analogous to $c : p$ (namely, proposition p is given in context c) in natural deduction, the operation of entering c can be seen as assuming p in c. Entering c, inferring another proposition q from p (as a result of noticing $p \rightarrow q$, say), and leaving c will let one assert $ist(c,q)$ in the outer context.

3.2. Logicists' works inspired by McCarthy

Guha (1991) models contexts with 'microtheories' and uses them in Cyc, a large-scale, highly modular commonsense reasoning program (Guha and Lenat 1994). Microtheories are theories of limited domains. They have two basic properties: (i) there is a set of axioms related to each microtheory, and (ii) there is a vocabulary which tells us the syntax and semantics of each predicate and each function specific to the microtheory. Different microtheories make different assumptions about the world. Similar to McCarthy's conception, they are interrelated via lifting rules stated in an outer context.

Shoham (1991) uses the alternative notation p^c to denote that assertion p holds in context c. Shoham's purpose is not really to offer a precise semantics for p^c. He is more interested in studying the interaction between modal operators (e.g., the knowledge operator K in the logic of knowledge) and context. His notion of contextual knowledge, denoted as $K^c p$ and meaning "p is known in context c," is a fitting example.

S. Buvač and Mason (1993)—and in a more recent work, S. Buvač, V. Buvač, and Mason (1995)—investigate the logical properties of contexts. They also use $ist(c,p)$ to denote context-dependent truth. Using this modality, they extend the classical propositional logic to what they call the 'propositional logic of context.' 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.

For Giunchiglia (1993), context is just a subset of facts from the knowledge base plus the reasoning machinery to compute with it. In formal terms, a context c_i is a triple $\langle \lambda_i, \alpha_i, \delta_i \rangle$ where λ_i is the language of the context, α_i is the set of axioms of the context, and δ_i is the inference mechanism of the context. Under this definition, there are bridge rules 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. Thus, the analogue of McCarthy's ist formula (asserted in c_0) becomes $\frac{<A,c>}{<ist(A,c),c_0>}$.

Attardi and Simi (1995) present a formalization of their notion of 'viewpoint,' meant for expressing varieties of relativized truth. Viewpoints denote sets of sentences which represent the axioms of a theory. The basic relation is $in('\sigma', vp)$. This says that σ is a sentence provable from viewpoint vp by means of natural deduction.

4. THE SITUATION THEORETIC APPROACH

Situation theory is a mathematical theory of information (Devlin 1991). Two of its primitive concepts are infons and situations. Infons are the basic units which embody discrete items of information. They are denoted as $\ll R, a_1, \ldots, a_n, i \gg$, where R is an n-place relation, $a_1, \ldots, a_n$ are objects appropriate for the respective argument places of R, and i is the polarity (1 if R holds, 0 if R does not hold).

A situation is a limited portion of the world (over some location and time), which can be picked out by a cognitive agent. It thus corresponds rather well to the intuitive meaning of 'situation' in English. For example, the sentence "I solved a puzzle during the invited talk of Cooper" describes an activity performed at a particular time and location, individuated as the situation 'the invited talk of Cooper.' Situations make certain infons factual. Using a notation deceptively hinting at first-order logic, s is said to support ι (symbolically, $s \models \iota$) provided that ι is an infon that is true of situation s.

Abstract situations are the mathematical (albeit ontologically impoverished) counterparts of real situations, and unlike the latter, are amenable to symbolic manipulation. Given a real situation s, the set $\{\iota \mid s \models \iota\}$ is taken to be the corresponding abstract situation. (This set will be nonwellfounded when s is a circular situation. However, this need not concern us in this paper.)

Let s be a given situation. Following the standard practice, we require the availability of some device for making reference to arbitrary objects of a given type[3], viz. parameters. If $\dot{x}$ is a parameter and I is a finite set of infons (involving $\dot{x}$), then there is a type $[\dot{x} \mid s \models \bigwedge_{\iota \in I} \iota]$. This is the type of those objects to which $\dot{x}$ may be anchored[4] in s, so that all the conditions in 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. Here $\dot{x}$ is the abstraction parameter and s is the 'grounding' situation.

In situation theory, the flow of information is realized via constraints. We denote a constraint as $S_1 \Rightarrow S_2$ (corresponding, in essence, to the infon $\ll involves, S_1, S_2, 1 \gg$, read '$S_1$ involves S_2'), where S_1 and S_2 are situation types. Cognitively, if this relation holds, then it is a fact that if S_1 is realized (i.e., there is a real situation s_1 of type S_1), then so is S_2 (i.e., there is a real situation s_2 of type S_2). For instance, the constraint $S_s \Rightarrow S_t$, where

$$S_s = [\dot{s} \mid \dot{s} \models \ll slaps, \dot{a}, \dot{b}, \dot{l}, \dot{t}, 1 \gg]$$
$$S_t = [\dot{s} \mid \dot{s} \models \ll touches, \dot{a}, \dot{b}, \dot{l}, \dot{t}, 1 \gg]$$

may be used to correctly infer that if Alice slapped Bob in a given situation (with spatio-temporal coordinates $\dot{l}$ and $\dot{t}$), then she touched him in that very same situation (Devlin 1991, p. 92).

4.1. Towards a formalization of context in situation theory

Following Barwise (1986), we will treat context as an amalgamation of a grounding situation and the rules which govern the relations within the context. Thus we will represent a context by a situation type that supports two kinds of infons: (i) factual infons to state facts, and (ii) constraints (which correspond to parametric conditionals) to capture the if-then relations holding within the context.

For example, let s be Sullivan's M.S. thesis presentation context at Drofnats University. Acker is Sullivan's advisor, and Leyner and Kraft are members of Sullivan's

[3] The basic types of situation theory include temporal locations, spatial locations, individuals, relations, situations, infons, etc. (Devlin 1991, p. 53).

[4] An 'anchor' is simply a function which assigns to each parameter in a set of parameters an object of a particular type.

jury. Accordingly, the context s supports infons such as the following:[5]

$$\begin{array}{l} \ll student, Sullivan, Drofnats\ University, 1 \gg \\ \ll msadvisor, Acker, Sullivan, 1 \gg \\ \ll msjurymember, Leyner, Sullivan, 1 \gg \\ \ll msjurymember, Kraft, Sullivan, 1 \gg \end{array}$$

Let us assume that at Drofnats, there exists an academic regulation valid for all thesis presentation contexts, given by the constraint C below:

$$\begin{array}{l} S_1 = [\dot{s} \mid \dot{s} \models \ll msadvisor, \dot{a}, \dot{b}, 1 \gg] \\ S_2 = [\dot{s} \mid \dot{s} \models \ll msjurymember, \dot{a}, \dot{b}, 1 \gg] \\ C = S_1 \Rightarrow S_2 \mid B \end{array}$$

C codifies the nonmonotonic rule that at Drofnats, advisors are usually jury members of their advisees. In terms of infons, the constraint C could be written as $\ll involves, S_1, S_2, B, 1 \gg$. The extra argument B in C is new and requires explanation (Barwise 1986; Devlin 1991). B is a set of background conditions under which C will convey information (rather than misinformation), and thus may profitably be employed by a cognitive agent 'attuned' to C. Basically, we have $\ll involves, S_1, S_2, 1 \gg$ as long as the background conditions in B are met.

One such condition in our example may be as follows:[6]

$$B \models \ll kinsman, \dot{a}, \dot{b}, 0 \gg$$

Then it is enough simply that the conditions in B obtain under the particular circumstances, i.e., $s \models \ll kinsman, Acker, Sullivan, 0 \gg$. Thus, using s as a grounding situation with the anchoring $f(\dot{a}) = Acker$ and $f(\dot{b}) = Sullivan$, we can infer, via C, that Acker must also be a member of Sullivan's jury.

After this example, let us review the desired properties of context, and check whether our proposal supports them.

4.2. Contexts vs. situations

During the review of McCarthy's work, we stated that contexts are first-class objects, so that one can use them in the same way as other objects. In our approach, we are modeling contexts with situation types, and situation types are situations which have some unbound parameters. Other than having unbound parameters, situation types are ordinary situations, and thus first-class objects of situation theory.

Richness of contexts was stated by McCarthy (1987, 1993) and Guha (1991). A rich object cannot be defined completely using extensional means. In situation theory, situations are, by definition, rich objects (Devlin 1991). Clearly, the richness of situations leads to the partiality of contexts, as McCarthy advocates.

Another aspect of the use of context is the flexibility of having private rules and presuppositions related to a particular point of view. In the logicist approach,

[5]For simplicity, spatio-temporal coordinates $\dot{l}$ and $\dot{t}$ are omitted throughout the example.

[6]For the sake of the argument, imagine another, bizarre academic regulation: if the advisor and the student are relatives—however distant—then the advisor cannot be in the jury.

presuppositions are represented with predicates that contain no variables and rules are usually represented with quantified logical implications, e.g.:

$$c : \; present(air)$$
$$c : \; \forall x \; bird(x) \rightarrow flies(x)$$

The first line states that air is present in the environment (a presupposition), and the second line states that if something is a bird then it flies (a default rule). The same capability is also available in our notion of context. We represent the facts related to a particular context with parameter-free infons supported by the situation type which corresponds to the context. The rules of the context are represented by constraints. Therefore, we can use C below to correspond to c:

$$S_1 = [\dot{s} \mid \dot{s} \models \ll bird, \dot{a}, 1 \gg]$$
$$S_2 = [\dot{s} \mid \dot{s} \models \ll flies, \dot{a}, 1 \gg]$$
$$B \models \ll present, air, 1 \gg \wedge \ll penguin, \dot{a}, 0 \gg \wedge \cdots$$
$$C = S_1 \Rightarrow S_2 \mid B$$

Here B is the set of conditions which render the default rule true. Barwise (1986, p. 124) points out to an intricate issue regarding such background conditions: "[T]he exact information content of a statement of a general conditional [is] highly context dependent, which seems right. However, it might appear to be too context dependent, since it could happen that the exact information content is not even determined by what the speaker knows, in that he or she might not know what the relevant conditions B are." In our model, the context representation is designed to supply just the adequate background information, e.g., context defines the domain of quantification. This property of context is due to its use as a grounding situation, so that in the binding of parameters, the only available objects are those available in the context.

5. REWORKING MCCARTHY'S LIFTING EXAMPLE

Lifting axioms are used to relate truth in one context to truth in another context. Since the vocabularies, languages, and assumptions of the source context and target context are usually different, these differences need to be addressed during the lifting. Lifting needs to be as 'meaning-preserving' as possible (Guha and Lenat 1994) and this makes it useful for NLP applications where preserving meaning is highly desirable, e.g., translation or natural language generation tasks. Consider the following simple scenario (Guha 1991, p. 35) which shows that a person interested in NLP should care about lifting:

> Fred is standing in front of Chris. There is a flower pot to the left of Chris. Fred says "I like that flower pot to your left." Let this statement be F_1 and the context in which this is uttered be C_1. Chris then moves so that the flower pot is to his right and tells Fred that he did not hear what Fred just said and asks him to say it again. Fred wants to convey the same message in this new context (C_2) but cannot use the same sentence (F_1). The sentence which states the same thing in this new context is "I like that flower pot to your right." Call this second sentence F_2. F_1 states in C_1 exactly what F_2 states in C_2. Given F_1, C_1, and C_2, the process of obtaining F_2 from F_1 is called lifting F_1 from C_1 to C_2.

In the remainder of this section, we will redo a lifting example due to McCarthy (1993). The example, just like Guha's, is conceptually trivial but is illustrative of the technicalities lifting poses in general.

McCarthy considers two contexts, namely, *Above-Theory (AT)* and c. AT is the context which expresses a static theory of the blocks world predicates *on* and *above*, cf. equations 1 and 2. In AT, the notion of situation—in the sense of situation calculus (McCarthy and Hayes 1969)—is not available.[7] However, we need to lift the results of AT to outer contexts that do involve situations—again in the sense of situation calculus—or times. The context c is such a context; it contains the theory of blocks world expressed using situation calculus, cf. equations 3, 4, and 5. For example, the predicate $on(x, y)$ (of AT) becomes $on(x, y, s)$ in c, where s denotes the situation in which $on(x, y)$ holds. We want to use AT in c. In other words, c needs to relate its predicates $on(x, y, s)$ and $above(x, y, s)$ to predicates $on(x, y)$ and $above(x, y)$ of AT. This is realized by $context - of(s)$, a function giving a context which depends on the situation parameter s. Equations 3 and 4 associate a context $context - of(s)$ with each situation s. Equation 5 is the major lifting rule, which asserts that the facts of AT all hold in the contexts associated with situations. The bottom line in McCarthy's example is, given $c_0 : \ ist(c, on(A, B, S_0))$ prove that $c_0 : \ ist(c, above(A, B, S_0))$. (As usual, c_0 is an outer context.)

Here are the axioms mentioned above, all in one place:

$$AT : \ \forall x \forall y \ on(x, y) \rightarrow above(x, y) \tag{1}$$
$$AT : \ \forall x \forall y \forall z \ above(x, y) \wedge above(y, z) \rightarrow above(x, z) \tag{2}$$
$$c : \ \forall x \forall y \forall s \ on(x, y, s) \leftrightarrow ist(context - of(s), on(x, y)) \tag{3}$$
$$c : \ \forall x \forall y \forall s \ above(x, y, s) \leftrightarrow ist(context - of(s), above(x, y)) \tag{4}$$
$$c : \ \forall p \forall s \ ist(AT, p) \rightarrow ist(context - of(s), p) \tag{5}$$

The proof of McCarthy proceeds as follows:

$$c_0 : \ ist(c, on(A, B, S_0)) \tag{6}$$
$$c : \ ist(context - of(S_0), on(A, B)) \tag{7}$$
$$context - of(S_0) : \ on(A, B) \tag{8}$$
$$c : \ ist(context - of(S_0), \forall x \forall y \ on(x, y) \rightarrow above(x, y)) \tag{9}$$
$$context - of(S_0) : \ \forall x \forall y \ on(x, y) \rightarrow above(x, y) \tag{10}$$
$$c : \ ist(context - of(S_0), above(A, B)) \tag{11}$$
$$c : \ above(A, B, S_0) \tag{12}$$

Briefly, equation 6 is the assumption given in the problem statement. Equation 7 is obtained from equations 3 and 6 by plugging A for x, B for y, and S_0 for s. Equation 8 obtained from equation 7 by entering $context - of(S_0)$. Equation 9 is the result of lifting equation 1 by the lifting rule (equation 5). Entering $context - of(S_0)$ we obtain equation 10. From equations 8 and 10, we obtain equation 11. Now using equations 4 and 11 we arrive at equation 12. The desired conclusion immediately follows from equation 12.

This proof can be visualized as in Fig. 1. In the figure, contexts are represented as Venn diagrams. Atomic formulas are represented with capital letters, and transfers

[7] A good justification for this is given in (McCarthy and Buvač 1994, p. 6): "In reasoning about the predicates themselves it is convenient not to make them depend on situations or on a time parameter."

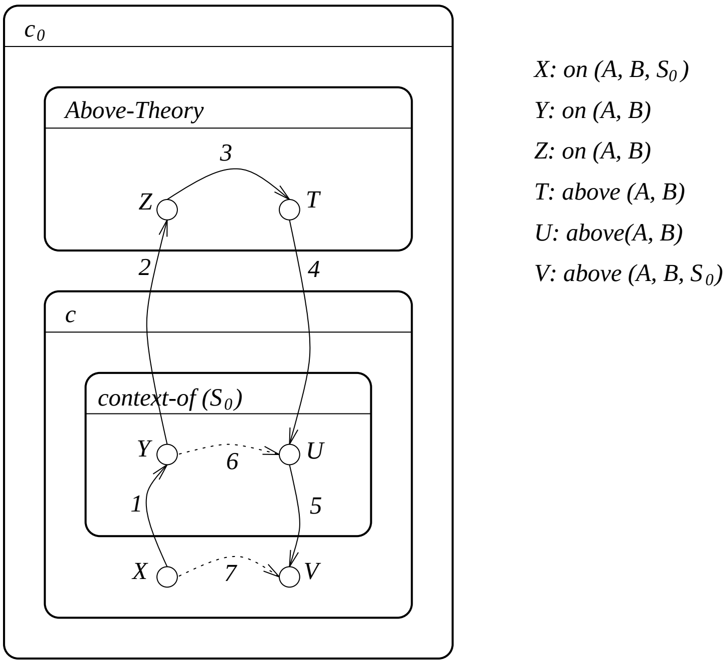

FIGURE 1. Diagram of McCarthy's proof.

between contexts are represented by arrows. We have labeled arrows in order to refer to the way the proof grows. Basically, McCarthy is drawing a virtual arrow from the atomic formula X to the atomic formula V. Since c has no rule to draw an arrow from X to V, he first creates $context - of(S_0)$ and draws an arrow to the atomic formula Y using equation 3. After Y, McCarthy lifts the implication of $above(x, y)$ from $on(x, y)$ (the arrow labeled with 3 in the figure) to $context - of(S_0)$, i.e., he forms the arrow labeled with 6. Then from Y, by tracing this arrow, he gets to U. From U, by leaving $context - of(S_0)$, he concludes with the desired formula V.

In the proof of McCarthy, it would be more natural to use the path 1-2-3-4-5. However, this path requires one more rule to transfer Y to Z (the arrow labeled with 2). In Attardi and Simi (1995), this is explicitly stated and a proof is carried out with the mentioned path. In the following reworking of McCarthy's example, we will also follow the path 1-2-3-4-5. But first some provisos:

1. In McCarthy's original example, *on* and *above* have different arities in different contexts. For instance, in AT *on*'s arity is two whereas in c its arity is three. However, in situation theory, we may refer to *on* and *above* in different contexts with different names (Devlin 1991, p. 115).[8]

[8] Thus, we are duplicating relations to facilitate different usages. While it does not simplify the analysis in the sequel, a more principled alternative—the one Devlin adopts—is to work with relations having a single, fixed number of argument places, but to allow the use of relations with unfilled argument roles.

2. Instead of $context - of(s)$ we will use $\ll context - of, \dot{s}, \dot{\sigma}, 1 \gg$ where $\dot{s}$ is a parameter of type situation (in situation theory) and $\dot{\sigma}$ is a parameter of type situation (in situation calculus). Thus, $\dot{s}$ is the context corresponding to $\dot{\sigma}$.
3. The contexts c_0, c, and AT of McCarthy will be represented with the situations (in situation theory) c_{c0}, c_c, and c_{AT}, respectively.
4. The background conditions B_{AT}, B_{cAT}, and B_{ATc}[9] will be shown below but will not be employed in the proof, since the original proof of McCarthy does not involve any nonmonotonic inference.[10]

The axioms of McCarthy will be captured by the following situation theoretic constructs (we do not need equation 5):

$$\begin{array}{l} S_{11} = [\dot{s} \mid \dot{s} \models \ll on_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ S_{12} = [\dot{s} \mid \dot{s} \models \ll above_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ c_{AT} \models \ll involves, S_{11}, S_{12}, B_{AT}, 1 \gg \end{array} \tag{13}$$

$$\begin{array}{l} S_{21} = [\dot{s} \mid \dot{s} \models \ll above_{AT}, \dot{x}, \dot{y}, 1 \gg \wedge \ll above_{AT}, \dot{y}, \dot{z}, 1 \gg] \\ S_{22} = [\dot{s} \mid \dot{s} \models \ll above_{AT}, \dot{x}, \dot{z}, 1 \gg] \\ c_{AT} \models \ll involves, S_{21}, S_{22}, B_{AT}, 1 \gg \end{array} \tag{14}$$

$$\begin{array}{l} S_{31} = [\dot{s} \mid \dot{s} \models \ll on_c, \dot{x}, \dot{y}, \dot{\sigma}, 1 \gg] \\ c_{AT} = [\dot{s} \mid \dot{s} \models \ll on_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ c_c \models \ll context - of, \dot{s}, \dot{\sigma}, 1 \gg \\ c_c \models \ll involves, S_{31}, c_{AT}, B_{cAT}, 1 \gg \\ c_c \models \ll involves, c_{AT}, S_{31}, B_{ATc}, 1 \gg \end{array} \tag{15}$$

$$\begin{array}{l} S_{41} = [\dot{s} \mid \dot{s} \models \ll above_c, \dot{x}, \dot{y}, \dot{\sigma}, 1 \gg] \\ c_{AT} = [\dot{s} \mid \dot{s} \models \ll above_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ c_c \models \ll context - of, \dot{s}, \dot{\sigma}, 1 \gg \\ c_c \models \ll involves, S_{41}, c_{AT}, B_{cAT}, 1 \gg \\ c_c \models \ll involves, c_{AT}, S_{41}, B_{ATc}, 1 \gg \end{array} \tag{16}$$

Initially, McCarthy has $c_0 : ist(c, on(A, B, S_0))$. The *on* relation can be represented with the infon $\iota_i = \ll on_c, A, B, S_0, 1 \gg$; this, in context c_{c0}, gives rise to $c_c \models \iota_i$. Finally, we need to conclude that in c_{c0}, we have $c_c \models \iota_f$, where $\iota_f = \ll above_c, A, B, S_0, 1 \gg$. This is what we will do now.

In our proof, we will first transfer the fact $c_c \models \iota_i$ to c_{AT}, then reason that *on* implies *above*, and finally carry this new fact to c_c. As noted earlier, this will be the path 1-2-3-4-5 in Fig. 1. Here is the proof briefly. Using the first constraint ($involves$) in equation block 15 with the anchoring $f_1(\dot{\sigma}) = S_0$, $f_1(\dot{s}) = s_0$, $f_1(\dot{x}) = A$, and $f_1(\dot{y}) = B$, we transfer $\ll on_c, A, B, S_0, 1 \gg$ from c_c to $\ll on_{AT}, A, B, 1 \gg$ in c_{AT}. This corresponds to the tracing of the arrows labeled 1 and 2 in Fig. 1. We will use the same anchoring function when we return to c_c. In c_{AT}, using the anchoring $f_2(\dot{x}) = A$ and $f_2(\dot{y}) = B$, and equation block 13, we get $\ll above_{AT}, A, B, 1 \gg$. This corresponds to the arrow labeled 3 in Fig. 1. After this implication of *above* from *on*, we should transfer the fact to c_c. This is done using the second constraint ($involves$) in equation block 16 with the anchoring f_1. The result is $\ll above_c, A, B, S_0, 1 \gg$.

[9] B_{AT} denotes the background conditions used in the constraints in AT, B_{cAT} the background conditions used in lifting from c to AT, and B_{ATc} the background conditions used in lifting from AT to c.

[10] Recall that when we use background conditions to implement nonmonotonicity, we are basically looking for the opposites of the background conditions to appear in the context. If we do not find any opposites in the context, we conclude that the background conditions are not being violated.

	Mc	Gu	Sh	Gi	Bu	At	Ba
Logic versus situation theory	Logic	Logic	Logic	Logic	Logic	Logic	S.T.
Modal logic	No	No	Yes	No	Yes	No	No
Natural deduction	Yes	Yes	No	Yes	Yes	Yes	No
Supports circularity	No	No	No	No	No	Yes	Yes
Partiality	Yes	Yes	No	Yes	Yes	?	Yes
Dynamic context	Yes	Yes	No	?	Yes	?	Yes
Natural language support	?	Yes	No	No	?	?	Yes

Mc $\Longrightarrow$ (McCarthy 1993; McCarthy and S. Buvač 1994)
Gu $\Longrightarrow$ (Guha 1991; Guha and Lenat 1994)
Sh $\Longrightarrow$ (Shoham 1991)
Gi $\Longrightarrow$ (Giunchiglia 1993)
Bu $\Longrightarrow$ (S. Buvač and Mason 1993; S. Buvač et al. 1995)
At $\Longrightarrow$ (Attardi and Simi 1995)
Ba $\Longrightarrow$ (Barwise 1986)

TABLE 1. Comparison of previous work and the situation theoretic approach. (? denotes lack of information regarding a particular issue, or an incomplete characterization.)

This completes the proof path 1-2-3-4-5. Using two anchoring functions (f_1 grounded at c_{c0} and f_2 grounded at c_{AT}), we have carried out the proof of McCarthy in our situation theoretic framework.

Let us emphasize the major idea in the above analysis. Basically, the logical reasoning of McCarthy is translated to an information-based reasoning, where the essential idea is to use the supports relation and constraints (with proper anchorings). It is noted that since a material equivalence, as in equations 3 and 4, can be written as a conjunction of two material implications, there are two symmetric constraints in each of equation blocks 15 and 16.

6. CONCLUSION

In the AI literature, there are a number of attempts towards a logical formalization of context. Our formal model of context differs from these in being stated in the framework of situation theory (Devlin 1991). The comparison of previous works and the situation theoretic approach is summarized in Table 1, where the first row categorizes the language of formalization. Since our work is essentially an application of Barwise's ideas (Barwise 1986), no attempt is made in Table 1 to add an extra column corresponding to our approach.

Compared to other approaches, our proposal has two notable properties:

1. *Dynamic contexts.* We might easily require the contents of a context change dynamically. We can add (delete) assumptions and rules into (from) a context. Having a dynamic notion of context is not a novel thing for the logicist, since he or she can always modify a theory. However, when we fortify our context with constraints whose background conditions are also dynamic, we get nonmonotonicity in the framework of situation theory.
2. *Natural language support.* Situation theory adopts a more natural outlook regarding natural language concepts (Barwise and Perry 1983). Thus, our approach might lead to simpler interfaces in NLP applications.

ACKNOWLEDGMENTS

We are deeply grateful to the guest editors of this special issue for their constructive criticism and expert technical advice. Thoughtful contributions of two anonymous referees have considerably improved the contents of the paper. Obviously, all the remaining inadequacies should be blamed on us.

REFERENCES

AKMAN, V., and M. SURAV. 1995. Contexts, oracles, and relevance. *In* Working Notes of the AAAI-95 Fall Symposium on Formalizing Context. *Edited by* S. Buvač. Technical Report FS-95-02. AAAI Press, Menlo Park, CA, pp. 23–30.

AKMAN, V., and M. SURAV. 1996. Steps toward formalizing context. AI Magazine, **17**(3): 55–72.

ASHER, R. E. (editor-in-chief) and J. M. Y. SIMPSON (coordinating editor). 1994. The encyclopedia of languages and linguistics. Volume 2. Pergamon Press, Oxford, UK.

ATTARDI, G., and M. SIMI. 1995. A formalization of viewpoints. Fundamenta Informaticae, **23**(3): 149–174.

BARWISE, J. 1986. Conditionals and conditional information. *In* On conditionals. *Edited by* E. C. Traugott, C. A. Ferguson, and J. S. Reilly. Cambridge University Press, Cambridge, UK, pp. 21–54.

BARWISE, J., and J. PERRY. 1983. Situations and attitudes. MIT Press, Cambridge, MA.

BLACKBURN, S. 1994. The Oxford dictionary of philosophy. Oxford University Press, Oxford, UK.

BUVAČ, S., and I. A. MASON. 1993. Propositional logic of context. Proceedings of the Eleventh National Conference on Artificial Intelligence, Washington, DC, pp. 412–419.

BUVAČ, S., V. BUVAČ, and I. A. MASON. 1995. Metamathematics of contexts. Fundamenta Informaticae, **23**(3): 263–301.

CRYSTAL, D. 1991. A dictionary of linguistics and phonetics. 3rd edition. Blackwell, Oxford, UK.

DEVLIN, K. 1991. Logic and information. Cambridge University Press, New York, NY.

FARWELL, D., and S. HELMREICH. 1995. Contextualizing natural language processing. Manuscript. Computing Research Laboratory, New Mexico State University, Las Cruces, NM.

GIUNCHIGLIA, F. 1993. Contextual reasoning. Epistemologia, **XVI**: 345–364.

GUHA, R. V. 1991. Contexts: a formalization and some applications. Technical report ACT-CYC-423-91. Microelectronics and Computer Technology Corporation, Austin, TX.

GUHA, R. V., and D. B. LENAT. 1994. Enabling agents to work together. Communications of the ACM, **37**(7): 127–142.

LEECH, G. 1981. Semantics: the study of meaning. 2nd edition. Penguin, Harmondsworth, UK.

LYONS, J. 1995. Linguistic semantics: an introduction. Cambridge University Press, Cambridge, UK.

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

———— 1993. Notes on formalizing context. Proceedings of the Thirteenth International Joint Conference in Artificial Intelligence (IJCAI-93), Chambery, France, pp. 555–560.

———— 1996. A logical AI approach to context. Manuscript. Computer Science Department, Stanford University, Stanford, CA.

MCCARTHY, J., and S. BUVAČ. 1994. Formalizing context (expanded notes). Technical note STAN-CS-TN-94-13. Computer Science Department, Stanford University, Stanford, CA.

MCCARTHY, J., and P. J. HAYES. 1969. Some philosophical problems from the standpoint of artificial intelligence. *In* Machine intelligence 4. *Edited by* B. Meltzer and D. Michie. Edinburgh University Press, Edinburgh, UK, pp. 463–502.

SHOHAM, Y. 1991. Varieties of context. *In* Artificial intelligence and mathematical theory of computation: papers in honor of John McCarthy. *Edited by* V. Lifschitz. Academic Press, Boston, MA, pp. 393–408.

SURAV, M., and V. AKMAN. 1995. Modeling context with situations. *In* Working Notes of the IJCAI-95 Workshop on Modeling Contextn in Knowledge Representation and Reasoning. *Edited by* P. Brézillon and S. Abu-Hakima. Technical report LAFORIA 95/11. Laboratoire Formes et Intelligence Artificielle, Université Paris VI, Paris, France, pp. 145–156.