CONTEXT IN ARTIFICIAL INTELLIGENCE: A FLEETING OVERVIEW

Varol Akman

Department of Computer Engineering, Bilkent University
Bilkent, Ankara 06533, Turkey

http://www.cs.bilkent.edu.tr/~akman
email: akman@cs.bilkent.edu.tr
phone: +90 312 290-1537
fax: +90 312 266-4047

Introduction

The notion of context arises in assorted areas of artificial intelligence (AI), including knowledge representation, natural language processing, intelligent information retrieval, etc. Although the term '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. In other words, it is used in an intuitive manner. In an influential paper, Clark and Carlson (1981) state that context has become a favourite word. They then complain that the denotation of the word has become murkier as its uses have been extended in many directions, making context some sort of 'conceptual garbage can.'

An example of how contexts may help in AI is found in a study (McCarthy, 1984) of MYCIN -- a program for advising physicians on treating bacterial infections of the blood and meningitis (Shortliffe, 1976). When MYCIN is told that the patient has Cholerae 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 diarrhoea. A contextual version of MYCIN, on the other hand, needs to add the context of a treatment and notice that any prescription must be made in the light of the fact that there is alarming dehydration.

The main motivation in AI for studying formal contexts is to resolve the 'problem of generality' (McCarthy, 1987). McCarthy believes that AI programs suffer from a lack of generality. A seemingly minor addition (as in the MYCIN example) to the predetermined possibilities that a program is required to handle would often necessitate a partial redesign of the program. Explicitly represented contexts might help because a program would then make its assertion about a certain context.

A more general objection to the implicit representation of context can be given as follows. Suppose we write longer (more complex) or shorter (simpler) 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 say "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 a rival (McCarthy, 1987: p. 1034):

> Consider axiomatizing *on* so as to draw appropriate consequences from the information expressed in the sentence, "The book is on the table." [An 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.

My aim in this paper is to offer a swift review of context in AI. I will first identify the role of context in various fields of AI. I will then consider the enterprise of formalising context. The focus of my discussion will be McCarthy's (1993) proposal – a landmark study that inspired almost every logical theory of context since then – and its close variants, studied by his students. Numerous recent, definitely more advanced and streamlined proposals will be omitted, essentially for two reasons. First, I think such a coverage would make this paper much longer. Second, I'd rather have the reader delve into the relevant papers and make up her own mind about their merits.[1]

Some Definitions

According to the Oxford English Dictionary, the term 'context' usually has two primary meanings:

1. the words around a word, phrase, statement, etc. often used to help explain (fix) the meaning;
2. the general conditions (circumstances) in which an event, action, etc. takes place.

Clearly, the first meaning is closely related to linguistic meaning and the linguists' use of the term, whereas the second meaning is the one which is closer to a desirable account of context in AI. This is in the spirit of the following observation (McCarthy, 1989: p. 180): "Almost all previous discussion of context has been in connection with 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."

In (Angeles, 1981), the same term is defined, reflecting the latter desideratum of McCarthy 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, judgements, and knowledge of that thing.

In 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.

Context in Natural Language

Context is a crucial factor in communication. Just consider the confusion that results from a 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, the discussants would realise this and try to give you a quick rundown of the conversations so far. This is essentially the view of Clark and Carlson (1981) 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. This is also known as the 'common ground.'

Leech (1981: p. 66) suggests another particularly attractive approach 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 disambiguating function (among others).[2] Consider the following conversation:

> A (a woman, talking to B): I am an investigator.

B (talking to C and referring to A): She is an investigator.
C (talking to A): So, you are an investigator.

In this segment, the word 'investigator' has a context-dependent meaning. The common ground of A, B, and C is used to select an appropriate meaning for this word. Similarly, the indexicals ('I' or 'she' or 'you') can be bound to the appropriate person (A) only by the help of context. For example, the sentences uttered by A and B have the same propositional content, and this we can say using some circumstantial information and conventions about discourse. To quote Recanati (1993: 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."

Another function of context arises when we deal with quantifiers in 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 in a universe, but only to those in a subset of the universe, determined by the contextual factors. Similarly for the meaning of 'many.' In an automobile factory, 100 automobiles might not qualify as many, but if a person owns 100 automobiles it surely counts as many.

Context might be used to fill the missing parameters in natural language utterances. Consider 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 Lewis in the Olympic Games, then the time and the place of the utterance are different from what we would get if we were watching him practice, prior to the Games, from our window.

Some relations stated in natural language necessarily need a context for disambiguation. Take the sentence

The Engineering Building is to the left of the Library.

In the context of the Bilkent Campus, if we are viewing these buildings from the Publishing House this sentence turns out to be true, but if we are located in the Tourism School it is false. More interestingly, if we are looking from the Faculty Club this sentence must be considered neither true nor false: the Library lies behind the Engineering Building.

It is seen that for natural languages a 'fleshing-out strategy' -- converting everything into decontextualized eternal sentences -- 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. Allen (1995, p. 465) explains: "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." The essential idea is that in discourse each new sentence should be interpreted (as a minimum) in the context provided by the sentences neighbouring it.

Context has long been a salient issue in social studies of language, viz. how human beings employ language to build the social and cultural organisations that they inhabit. Lyons (1995: p. 292) finds this 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." Goodwin and Duranti (1992: p. 32) judge context as a basic concept in ethnographically oriented studies of language use. They claim that this notion "stands at the cutting edge of much contemporary research into the relationship between language, culture, and social organisation, as well as into the study of how language is structured in the way it is." Akman (2000) offers a wide survey of context just from this perspective.

Context in Categorisation

Categorisation is one of the basic mental processes in cognition (Rosch, 1978). Human beings can categorise various types of objects, events, and states of affairs, where categorisations depend on the circumstance. Consider the following scenario:

> In Springfield, there are three barbers working for money, and a man who 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 (1992) use natural regularities to study the role of context in categorisation. Thus, "Swans are white" 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 and correct reference. And they are fallible since they are needed to account for misinterpretation 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. The 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. Natural regularities intuitively correspond to material implication in logic, and the effect of contextual factors is similar to the effect of nonmonotonicity. On the other hand, implication and nonmonotonicity are usually studied in logic 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 on and so forth' clause to cover the unexpected contextual factors, for it is simply impossible to state all of the relevant ones.

Context in Information Retrieval

A formal notion of context might be useful in information retrieval (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 'precision and recall' may be enhanced.

Traditional methods of IR use statistical methods to find the similarities between documents and the relevance of documents to a 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 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 that are based on the frequency of occurrence. Lately, psychological, epistemic, and semantical considerations are beginning to flourish.

In a pioneering study of relevance (Sperber and Wilson, 1986), this important notion is regarded as the psychological pertinence of a proposition to a context. The 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. Moreover, intuitions of relevance are relative to contexts, and there is no way of controlling exactly which context someone has in mind at a given moment.

Despite these difficulties, Sperber and Wilson 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:

1. Contextual implication: A new assumption can be used together with the existing rules in the context to generate new assumptions;
2. Strengthening: A new assumption can strengthen some of the existing assumptions;
3. 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 encyclopaedic information involved, and so on. Finally, they propose their celebrated maxim:

> An assumption is relevant in a context to the extent that its contextual effects in this context are large. (Conversely, an assumption is irrelevant in a context to the extent that the effort required to process it in this context is large.)

Harter (1992) uses this to interpret psychological relevance in relation to IR. According to him, reading a new bibliographic citation[3] 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.

Context in Knowledge Representation

When we state something we do so in a context. For example, 37 degrees centigrade would appear 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 point is that if we are to reason in a common sense way, we have to use certain contexts.

Being the largest commonsense knowledge building attempt, CYC (Lenat, 1995) has crucial pointers on reasoning with an explicit notion of context:

- The language (i.e., the predicates, functions, and categories) used for representation should be appropriate for their intended domain. For example, of the two renowned medical diagnosis programs ONCOCIN has some concept of time whereas MYCIN does not.
- The application area (context) determines the granularity and accuracy of the theory.

CYC researchers identify two approaches to building large commonsense knowledge bases, and reasoning with them. The straightforward way is to introduce an expressive and powerful vocabulary. This approach increases the complexity of the problem, since using a rich vocabulary causes difficulties in truth maintenance, and leads to large search spaces.

The second way (Guha, 1991) 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 relative decontextualization.

In CYC, Guha employs contexts regularly:

- A general theory of some topic: Contexts can encapsulate a theory of TV repair, a theory of weather in Rome, a theory of what to look for when buying cellular phones, etc. Such contexts are called 'micro-theories' (Guha, 1991). 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 context. We collect all related assumptions, rules, etc. in this problem-solving context, and can process a group of related queries in a relatively small search space.
- Context-dependent representation of utterances: We can use anaphoric and indefinite statements without completely decontextualizing them.

We'll return to Guha's approach in the sequel.

McCarthy on Contexts

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 denotes relativized-truth-within-a-context via a special predicate *ist(p,c)*. This states that proposition *p* holds (is true) in context *c*.

McCarthy (1993) gives some reasons for introducing the formal notion of context.

First, the use of context allows simple axiomatizations. He exemplifies this by stating that axioms for static blocks world situations can be 'lifted' 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.[4]

Third, contexts provide 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. However, this brings about two problems:

1. *When* to transcend a context? Either the system must be capable enough to do so or we must instruct it when to transcend one or more levels up.
2. *Where* to transcend? This can be answered if we are prepared to accept that formulas are always considered to be asserted within a context.

Given *ist(c,p)*, the main formulas in a knowledge base are sentences of the following kind:

c': ist(c,p)

Briefly, *p* is true in a context *c*, and this itself is asserted in an outer context *c'*. To give an example,

c': ist(context-of("Sherlock Holmes stories"),"Holmes is a detective")

asserts that it is true in the context of Sherlock Holmes stories that Holmes is a detective. Here, *c'* is the 'outer context.'

Two primary features of context are the following:

- Contexts are abstract objects: Many contexts will be 'rich objects.' A system may be given some facts about such an object but never the complete description, e.g., the situations in Situation Calculus. Some contexts will not be as rich and might be fully described, e.g., simple micro-theories (Guha, 1991).
- Contexts are 'first-class' objects: We can use contexts in our formulas in the same way we use other objects.

There are some relations working between contexts. The most notable one is ≤. This defines a partial ordering over contexts; $c \leq c'$ means that the second context c' contains all the information of the first context c and possibly more. Using ≤, we can 'lift' (to be explained in the sequel) a fact from a context to one of its super-contexts using the following nonmonotonic rule:

$$(c)(c')(p)\ c \leq c' \cdot ist(c,p) \cdot \sim ab1(c,c',p) \rightarrow ist(c',p)$$

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

$$(c)(c')(p)\ c \leq c' \cdot ist(c',p) \cdot \sim ab2(c,c',p) \rightarrow ist(c,p)$$

The difference between the abnormality relations is crucial: *ab1* represents the abnormality in generalising to a super-context, whereas *ab2* corresponds to the abnormality in specialising to a sub-context.

Here are some functions on contexts that we might want to define:

- *value(c,t)* is a function which returns the value of term *t* in context *c*. For example, *value(context-of("Sherlock Holmes stories"),"number of wives of Holmes")=0*, meaning Holmes has no wife in the context of Sherlock Holmes stories.
- *specialise-time(c,t)* is a function which returns a context related to *c* in which the time is specialised to the value *t*. For example, *ist(specialise-time(c,5/17/2002),at(VA,Bilkent))* may be asserted in an outer context to state that at time *5/17/2002* in context *c*, the author is at Bilkent University.[5]

McCarthy's theory of context can be used to model inference in the style of deduction. Thus, *assuming(c,p)* is another context like context *c* in which proposition *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 (McCarthy and Buvac, 1998):

- 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 (1987) suggested -- the operations of 'entering' and 'leaving' a context might be useful and shorten the proofs involving contexts. In this case, *ist(c,p)* will be analogous to $c \rightarrow p$, and the operation of entering *c* can be regarded as *assuming(c,p)*. Then, entering context *c* and inferring proposition *p* will be equivalent to asserting *ist(c,p)* in the outer context.

Lifting

By 'lifting' a predicate from one context to another, we mean transferring the predicate to the other context with appropriate changes. Here are some of the things we can do with lifting:

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

  $$c': (x)(y)\ ist(above\text{-}theory,on(x,y)) \rightarrow ist(blocks\text{-}world,on(x,y))$$

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

  $$c': (x)(y)(s)\ ist(above\text{-}theory,on(x,y)) \rightarrow ist(context\text{-}of(s),on(x,y))$$

  where *context-of* is the function (seen earlier) 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. This is too obvious to warrant an example.

Other Issues

Relative decontextualization is an approach 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. McCarthy thinks 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. Sentences in this context are relatively eternal. (A similar idea is used in CYC.)

Another place where context might be useful is the representation of 'mental states.' McCarthy (1993) proposes a scheme, in which mental states can be thought of as outer sentences; for example,

believe(paolo,publisher(thisbook)=mcgraw-hill,because…)

where ellipsis denotes the reasons for Paolo's belief that this book is published by McGraw-Hill. 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 – when no belief revision is required -- we simply enter the related context and assert them.

Guha on Contexts

Guha (1991) finds an essential use for formal contexts in implementing his so-called micro-theories. Micro-theories are formal accounts 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:

- 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 might 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 relativized to location. In an agency, the first thing to do may be to take away a file of sensitive documents, whereas, in a house, children must be saved first.

Lifting rules might be used to transfer facts from a source context to a 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 nonmonotonic and Guha uses default reasoning in stating them. 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 quite a 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.

S. Buvac et al. on Contexts

S. Buvac and Mason (1993) -- and later, S. Buvac, V. Buvac, and Mason (1995) -- approach context from a mathematical viewpoint. They investigate the logical properties of contexts. They use *ist(c,p)* to denote context-dependent truth and extend the classical propositional logic to a 'propositional logic of context.'[6] 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. Buvac and Mason discuss the syntax and semantics of a general propositional language of context, and give a Hilbert-style proof system for this language. They also provide a model theory for contexts and give soundness and completeness proofs. Their system has the following two features:

1. A context is modelled by a set of partial truth assignments that describe the possible states of affairs in that context. The modality *ist* is interpreted as validity: *ist(c,p)* is true if and only if *p* is true in all the truth assignments associated with *c*.
2. The nature of particular contexts is itself context-dependent. A good example 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. Buvac and Mason's terminology such a property is known as 'non-flatness.' The acceptance of a sequence of contexts respects the intuition that what holds in a particular context can depend on how that context is reached.

S. Buvac 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. They then show that introducing the outermost context only simplifies the way they are dealing with non-flatness.

Conclusion

I have reviewed the use and benefits of contexts for AI. McCarthy's landmark proposal and its close relatives were covered. While the idea of formalising context seems to have caught on and produced good theoretical outcomes, the area of innovative and convincing[7] applications remains relatively unexplored. This is evidently the point to which further research on context should converge.[8]

Acknowledgements

Many thanks to Carlo Penco for his invaluable support and congeniality. This chapter is a compressed and revised version of an earlier paper (Akman and Surav, 1996). I am grateful to the Editor-in-Chief of AI Magazine and AAAI Press for permission to reuse, with some changes, material from that work.

References

Akman, V. 2000. Rethinking context as a social construct. Journal of Pragmatics 32(6): 743-759.

Akman, V., Surav, M. 1996. Steps toward formalizing context. AI Magazine 17(3): 55-72.

Akman, V., Bouquet, P., Thomason, R., Young, R. A., eds. 2001. Modeling and Using Context (Proceedings of CONTEXT'01), Berlin: Springer.

Allen, J. 1995. Natural Language Understanding. Redwood City, California: Benjamin/Cummings.

Angeles, P. A. 1981. Dictionary of Philosophy. New York: Harper and Row.

Barwise, J., Seligman, J. 1992. The rights and wrongs of natural regularity. In Philosophical Perspectives 8: Logic and Language, J. Tomberlin, ed. Atascadero, California: Ridgeview Press.

Bonzon, P., Cavalcanti, M., Nossum, R., eds. 2000. Formal Aspects of Context (Selected Papers from CONTEXT'97). Dordrecht, The Netherlands: Kluwer.

Bouquet, P., Serafini, L., Brezillon, P., Benerecetti, M., Castellani, F., eds. 1999. Modeling and Using Context (Proceedings of CONTEXT'99), Berlin: Springer.

Buvac, S. 1996. Quantificational logic of context. In Proceedings of the Thirteenth National Conference on Artificial Intelligence. Menlo Park, California: AAAI Press.

Buvac, S., Buvac, V., Mason, I. A. 1995. Metamathematics of contexts. Fundamenta Informaticae 23(3): 263-301.

Buvac, S., Mason, I. A. 1993. Propositional logic of context. In Proceedings of the Eleventh National Conference on Artificial Intelligence, 412-419. Menlo Park, California: AAAI Press.

Clark, H. H., Carlson, T. B. 1981. Context for comprehension. In Attention and Performance IX, J. Long and A. Baddeley, eds., 313-330. Hillsdale, N.J.: Lawrence Erlbaum.

Goodwin, C., Duranti, A. 1992. Rethinking context: an introduction. In Rethinking Context: Language as an Interactive Phenomenon, A. Duranti and C. Goodwin, eds., 1-42. New York: Cambridge University Press.

Guha, R. V. 1991. Contexts: a formalization and some applications. Ph.D. thesis, Computer Science Department, Stanford University, Stanford, California.

Harter, S. P. 1992. Psychological relevance and information science. Journal of the American Society for Information Science 43: 602-615.

Leech, G. 1981. Semantics: The Study of Meaning. Harmondsworth, UK: Penguin.

Lenat, D. B. 1995. CYC: a large-scale investment in knowledge infrastructure. Communications of the ACM 38(11): 33-38.

Lyons, J. 1995. Linguistic Semantics: An Introduction. Cambridge, UK: Cambridge University Press.

McCarthy, J., Buvac, S. 1998. Formalizing context (expanded notes). In Computing Natural Language, A. Aliseda, R. J. van Glabbeek, and D. Westerstahl, eds., 13-50. Stanford, California: CSLI Publications.

McCarthy, J. 1993. Notes on formalizing context. In Proceedings of the Thirteenth International Joint Conference in Artificial Intelligence, 555-560. San Mateo, California: Morgan Kaufmann.

McCarthy, J. 1989. Artificial intelligence, logic, and formalizing common sense. In Philosophical Logic and Artificial Intelligence, R. H. Thomason, ed., 161-190. Dordrecht, The Netherlands: Kluwer.

McCarthy, J. 1987. Generality in artificial intelligence. Communications of the ACM 30(12): 1030-1035.

McCarthy, J. 1984. Some expert systems need common sense. In Computer Culture: The Scientific, Intellectual, and Social Impact of the Computer, H. Pagels, ed., 129-137. New York: Annals of the New York Academy of Sciences.

Recanati, F. 1993. Direct Reference: From Language to Thought. Oxford: Blackwell Publishers.

Rosch, E. 1978. Principles of categorization. In Cognition and Categorization, E. Rosch and B. B. Lloyd, eds. Hillsdale, N.J.: Lawrence Erlbaum.

Shortliffe, E. 1976. MYCIN: Computer-Based Medical Consultations. New York: Elsevier.

Sperber, D., Wilson, D. 1986. Relevance: Communication and Cognition. Oxford: Basil Blackwell.

---

[1] Useful resources in this regard are (Bouquet et al., 1999), (Bonzon et al., 2000), and (Akman et al., 2001).

[2] Leech (1981: p. 67): "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)."

[3] The setting here is that of a user, accepting or rejecting a bibliographic document retrieved by a library information system. (An analogous setting can be specified for a user searching the Web via search engines.)

[4] McCarthy's approach corresponds to the use of local variables 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.

[5] Instead of specialising on time, we can also specialise on location, speaker, situation, subject matter, and so on.

[6] The quantificational logic of context is treated in (Buvac, 1996) .

[7] This amounts to proving that the contextual viewpoint has decisive advantages compared to another approach which does not use contexts.

[8] The reader is referred to (Akman et al., 2001) for a glimpse of the state-of-the-art.