A LOGICAL AI APPROACH TO CONTEXT

John McCarthy
Computer Science Department
Stanford University
Stanford, CA 94305
jmc@cs.stanford.edu
http://www-formal.stanford.edu/jmc/

1996 Feb 6, 12:09 p.m.

Abstract

Logical AI develops computer programs that represent what they know about the world primarily by logical formulas and decide what to do primarily by logical reasoning—including nonmonotonic logical reasoning. It is convenient to use logical sentences and terms whose meaning depends on context. The reasons for this are similar to what causes human language to use context dependent meanings. This note gives elements of some of the formalisms to which we have been led. Fuller treatments are in [McC93], [Guh91] and [MB94] and the references cited in the Web page [Buv95]. The first main idea is to make contexts first class objects in the logic and use the formula $ist(c, p)$ to assert that the *proposition* p is true in the *context* c. A second idea is to formalize how propositions true in one context transform when they are moved to different but related contexts. An ability to *transcend* the outermost context is needed to give computer programs the ability to reason about the totality of all they have thought about so far [McC96].

1 Introduction

As requested by Johan van Benthem, this is a brief introduction to the logical formalism for context being explored by John McCarthy and Saša Buvač at

Stanford University. It is motivated by the need to use contexts as first order objects for artificial intelligence. I hope the description is suitable for comparison with other approaches to context that often have other motivations.

2 Features of the Formalism

Here are some features of our formalizations.

1. We offer no *definition* of context. There are mathematical context structures of different properties, some of which are useful. Asking what a context is is like asking what a group element is. See section 4 for more on this.

2. Sentences about propositions and contexts are built up from a formula $ist(c, p)$ which is to be understood as asserting that the *proposition* p is true in the *context* c. When we have entered the context c, we can write
$$c: \qquad p. \tag{1}$$

3. Once a program has inferred a sentence q from p, it can *leave* the context c and have $ist(c, q)$. This generalizes natural deduction.

4. Reasoning and communicating in context permits taking only limited phenomena into account. Treating contexts as objects permits stating the limitations explicitly within the formalism.

5. Statements about contexts are themselves in contexts.

6. There is no universal context. This is a fact of epistemology (both of the physical world and the mathematical world). It is always possible to generalize the concepts one has used up to the present. Attempts at ultimate definitions always fail—and usually in uninteresting ways. Humans and machines must start at middle levels of the conceptual world and both specialize and generalize.

7. We can deal with this phenomenon in our formalism by ensuring that it is always possible to *transcend* the outermost context used so far. Thus a robot designed in this way is not stuck with the concepts it has been given.

8. Because of the possibility of transcendence, the use of contexts as objects is not just a matter of efficiency. Any given set of sentences including contexts can always be *flattened* (at the cost of lengthening) to eliminate explicit contexts. However, the resulting flat theory can no longer be transcended within the formalism, because it is not an object that can be referred to as a whole.

9. There is often a theory associated with a context—the set of sentences true in the context. However, two contexts with the same theory need not be the same, because they may have different relations with other contexts. Not all useful contexts will be closed under logical inference.

10. We advocate using *propositions* as discussed in [McC79] for the objects true in contexts rather than logical or natural language sentences. This has the advantage that the set of propositions true in a context may be finite when the set of sentences that can express these propositions will be infinite. However, our present applications of context would work equally well if sentences were used. Buvač and Mason [BBM95] treat $ist(c, p)$ as a modal logic formula in a propositional theory.

11. Besides the truth of propositions in contexts, we consider the value $value(c, exp)$ of a term exp representing an *individual concept* in a context c as discussed in [McC79]. This presents problems beyond those presented by propositions, because in general the space of values of individual concepts will depend on some outer context.

3 Applications

Here are some applications of the logical theory of contexts.

1. Conventional linguistic applications like the referents of pronouns can be treated using contexts as objects, but formalized contexts are also useful for more complex anaphora. For example, we need to relate the surgeon's "Scalpel" to the sentence "Please hand me a number 3 scalpel". See [Buv96]. These applications require associating contexts with sentences or parts of sentences.

2. Defining a theory in a narrow context in a way that permits it to be *lifted* to a richer outer context and applied. [McC93] discusses lifting a

simple theory of $above(x, y)$ as the transitive closure of $on(x, y)$ to an outer situation calculus context that uses $on(x, y, s)$ and $above(x, y, s)$. A key formula of that paper is

$$c: \qquad (\forall xys)(on(x, y, s) \equiv ist(context\text{-}of\text{-}situation(s), on(x, y))), \tag{2}$$

which relates the three argument situation calculus predicate $on(x, y, s)$ and the two element predicate $on(x, y)$ of the specialized theory of *on* and *above*. The use of contexts to implement "microtheories" in Cyc is described in [Guh91]. This allowed people entering knowledge about some phenomenon, e.g. automobiles, to do it in a limited context, but leave open the ability to use the knowledge in a larger context.

3. Defining a narrow context for a problem and importing facts that permit the problem to be solved by considering only a small set of possibilities. For example, in formulating the missionaries and cannibals problem a person or program must take a number of common sense facts into account, but ends up with a 32 state space, because all that is relevant in this context is the numbers of missionaries, cannibals and boats on each bank of the river.

4. Relating databases with different conventions [MB94]. Imagine that the Airforce and the General Electric Company have databases both of which include prices for the jet engines that the company sells the Airforce. However, suppose the databases don't agree on what the price covers, e.g. spare parts. We use one context c_{AF} for the Air Force database, another c_{GE} for the GE database, and a third context $c0$ that needs to relate information from both. *Lifting* formulas in the context true in $c0$ relate information in the different databases to the context in which reasoning is done, , e.g. they tell about the relation of the prices listed in c_{AF} and c_{GE} to the inclusion or not of spare parts.

5. Buvač and McCarthy have also discussed using context to combine aspects of plans generated by different planners not originally designed to work together—or plans originally intended to work together but which have drifted apart in the course of separate development.

4 Desiderata for a Mathematical Logic of Context

The simplest approach to a logic of context is to treat $ist(c, p)$ as a modal operator with p quantifier free. Saša Buvač and Ian Mason [BBM95] did this. However, the applications to natural language, to databases and to formalizing common sense knowledge and reasoning require a lot more. Here are some desiderata for a formal theory.[1]

- $truths(c)$ is the set of p such that $ist(c, p)$. In some formalizations it will be a first class object. In any case we can think about it in the metatheory.
- The simplest possibility for $truths(c)$ for a particular context c is that it is an arbitrary set of propositions, i.e. not required to be closed under some logical operations.
- The second possibility is that $truths(c)$ is closed under deduction in some logical system—perhaps the theory of contexts.
- $truths(c)$ may be the set of propositions true about some subject matter. We can assert propositions about this set of proposition without knowing what sentences are in it.
- Associated with at least some contexts is a domain $domain(c)$. As with $truths(c)$, $domain(c)$ may be an object, presumably in a higher level context, or it may be only in the metalanguage.

The variety of potential applications of contexts as objects suggests looking at contexts as mathematics looks at group elements. Groups were first identified as sets of transformations closed under certain operations. However, it was noticed that the integers with addition as an operation, the non-zero rationals with multiplication as an operation and many others had the same algebraic property. This motivated the definition of abstract group around the turn of the century. In such a theory, formulas express relations among contexts would be primary rather than the propositions true in the contexts. Thus the theory would emphasize $specializes(c1, c2, time)$ rather than $ist(c, p)$.

[1] Just so Johan doesn't get off too easily in keeping his promise to make one.

5 Remarks

Johan van Benthem asked for the following in soliciting this essay and John Perry's.

> *My proposal is the following. I would like to invite the two Johns to send me a rough outline of their contribution. It would be good if you could bring out (1) what the notion of context* ***is*** *and what it* ***does*** *according to you: in both cases, I think you want it to achieve 'efficiency' and 'portability' of information, (2) what is involved in the dynamics of* ***changing contexts****, perhaps with attendant changes in linguistic formulation (add or drop variables, etcetera). I would then like to comment on this, adding some thoughts on possible logical formalizations, emphasizing the interplay between what is said in a formula and what remains implicit in the models where it gets evaluated.*

I have rejected the idea of defining what a context **is**, but I hope I have given some idea of what they do. The example relating the three argument *on* and the two argument *on* should provide a basis for comments. In the formulation of the ideas, the ability to combine formulas arising in different contexts has been more important than computational efficiency.

[McC93] and [MB94] have additional references. Also Saša Buvač has several other papers on context on his Web page http://www-formal.stanford.edu/buvac/.

References

[BBM95] Saša Buvač, Vanja Buvač, and Ian A. Mason. Metamathematics of contexts. *Fundamenta Informaticae*, 23(3), 1995.

[Buv95] Saša Buvač. Saša buvač's web page, 1995. http://www-formal.stanford.edu/buvac/.

[Buv96] Saša Buvač. Resolving lexical ambiguity using a formal theory of context. In *Semantic Ambiguity and Underspecification*. CSLI Lecture Notes, Center for Studies in Language and Information, Stanford, CA, 1996.

[Guh91] R. V. Guha. *Contexts: A Formalization and Some Applications*. PhD thesis, Stanford University, 1991. Also published as technical report STAN-CS-91-1399-Thesis, and MCC Technical Report Number ACT-CYC-423-91.

[MB94] John McCarthy and Saša Buvač. Formalizing Context (Expanded Notes). Technical Note STAN-CS-TN-94-13, Stanford University, 1994.

[McC79] John McCarthy. First order theories of individual concepts and propositions. In Donald Michie, editor, *Machine Intelligence*, volume 9. Edinburgh University Press, Edinburgh, 1979. Reprinted in [McC90].

[McC90] John McCarthy. *Formalizing Common Sense: Papers by John McCarthy*. Ablex Publishing Corporation, 355 Chestnut Street, Norwood, NJ 07648, 1990.

[McC93] John McCarthy. Notes on formalizing context. In *IJCAI-93*, 1993. Available on http://www-formal.stanford.edu/jmc/.

[McC96] John McCarthy. Making robots conscious of their mental states. In Stephen Muggleton, editor, *Machine Intelligence 15*. Oxford University Press, 1996. to appear, available on http://www-formal.stanford.edu/jmc/.

/@steam.stanford.edu:/u/jmc/f95/context.tex: begun 1995 Sep 22, latexed 1996 Feb 6 at 12:09 p.m.