文化論集第24号
2004年3月

Logics for Context in Natural Language

Mayumi Masuko

1. Introduction

Any theory of natural language semantics must have some mechanism to treat context. This is particularly important when dealing with referential expressions such as pronouns and demonstratives.

In Japanese, probably the most difficult and contentious issue is how to treat the so-called reflexive, *jibun*. This is mainly because the referential range of *jibun* cannot be accounted for simply by syntactic constraints and that adding some semantic/pragmatic factors to a basically syntactic explanation is insufficient. What is required is some mechanism that can represent "context" properly and can explain how plausibility is computed.

For this purpose, attempts by Artificial Intelligence researchers to formalise the notion of context appear useful. There are several different versions of logics of context and most of them are propositional. In this paper, representative theories will be examined and compared.

2. McCarthy's Logic of Context

The most widely used framework is the one proposed by John McCarthy (e.g. McCarthy, 1983 and 1996; McCarthy and Buvaç, 1997). This work, however, as its main proponent admits, remains "incomplete and tentative" (McCarthy and Buvaç, 1997, p.14; see de Paiva (2003) for more problems related to formalisation).

In this framework, contexts are regarded as abstract, formal objects: they are also said to be "*rich* objects, like situations in situation calculus" (McCarthy and Buvaç, 1997, 15) but no explication is offered as to what is meant by this. Hence, I shall only present how this logic is said to work, concentrating on notation.

Firstly, *ist* (c, p) means 'the proposition p is true in the context c'. *value* (c, e), on the other, designates the value of a term e in the context c. Thus, *value* $(c, x) = y \equiv (\forall z)\ y = z \equiv$ *ist* $(c, x = z)$.

One important notion in this theory is *transcending* contexts, which makes it possible to account for more than one subjects. This is represented as c': *ist* (c, p): i.e. the proposition p is in the context c, and this is asserted in an outer context c'. In addition, in order to explain reference relations in multiple contexts, it is necessary to *enter* and *exit* contexts. The outer context is *c0*, and if c0: *ist* (c, p), by entering the context c, it can be inferred that c: p. And by reversing the process, if we have c: p, we can infer c0: *ist* (c, p) by EXITing the context c.

I shall merely point out at this stage that the notion of 'outer context' (i.e. c0 in McCarthy and Buvaç (1997)) seems an appropriate means to represent the possibility of *jibun* referring to the speaker of a given utterance. Another useful feature is its capacity of allowing to use different vocabularies in different contexts; we shall come back to this point shortly.

3. Attardi & Simi (1993)

Attardi and Simi (1993) explicate viewpoints using a reflective first order logic that is proved to be consistent. In this framework, a viewpoint is seen as a set of sentences that represent the assumptions of a theory. Thus, in their notation, in ('A', vp) means that a sentence A is entailed by the assumptions denoted by a viewpoint expression vp. Belief, reflection, truth and knowledge (= true belief) are defined as follows (Attardi and Simi, 1993, p.15f.):

BELIEF	$\mathsf{Bel}\ (g, A) = \mathsf{in}\ (A, vp\ (g))$ where g is an agent.
Reflection	$\mathsf{in}\ (A, vp) \Rightarrow (\mathrm{vp} \Rightarrow A)$
TRUTH:	$\mathsf{True}\ (A) = \mathsf{in}\ (A, \mathsf{RW})$ where **RW** is a special theory called **Real World** that represents the real world we live in. Thus, $\mathsf{in}\ (\mathsf{in}\ (A, \mathsf{RW}), vp) \Leftrightarrow \mathsf{in}\ (A, vp)$
KNOWLEDGE:	$\mathsf{K}(g, A) = \mathsf{Bel}\ (g, A) \wedge \mathsf{True}\ (A) = \mathsf{in}\ (A, vp\ (g)) \wedge \mathsf{in}\ (A, \mathsf{RW})$. Thus, $\mathsf{K}\ (g, A) \Rightarrow A$

It should be clear from the above that truth is relative in this theory. Provability in a viewpoint is called holding in a situation, which is represented as: Hold(A, s) = in (A, vp (s)), where s is a situation, and a viewpoint vp (s), which is a set of basic facts which define the situation.

With this mechanism, Attardi and Simi can represent contexts with viewpoint as $ist\ (c, p) = \mathrm{in}\ (p, c)$. This, however, does not allow differences in vocabularies in different contexts, which is allowed in McCarthy's theory.

4. Ghidini and Giunchiglia (2001 & 2002)

Ghidini and Giunchiglia (2001; 2002) advocate a framework called Local Models Semantics (LMS). In this system, a context is seen as a partial and approximate representation of the world from some agent's perspective: i.e. it does not belong to the real world, as it were, but represents the world from some individual's viewpoint. Thus, reasoning is partial in a sense that it only involves a subset of the individual's knowledge and also that not all inference patterns will be used. This does not mean, however, that different contexts are unrelated. Ghidini and Giunchglia's argument is that such relationships between different contexts are deemed to be partial and we cannot fully 'translate' one context into another: a single representation of the real world is in principle not feasible.

Such intuitions are stated as two principles:

Principle 1 (of Locality): reasoning uses only part of what is potentially available (e.g., what is known, the available inference procedures). The part being used while reasoning is what we call *context*.

Principle 2 (of Compatibility): there is compatibility among the reasoning performed in different contexts.

(Ghidini and Giunchiglia, 2000 and 2001, p.2)

More formally, $\{L_i\}_{i \in I}$ is defined as a family of languages defined over a set of indexes *I*. In order to pair local models into a single uniform structure, a notion of a compatibility sequence **c** is defined as $\langle \mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_i \ldots \rangle$. A model in this framework is a compatibility relation **C** which has the following characteristics:

1. $\mathbf{C} \neq 0$
2. $\langle 0, 0, \ldots, 0, \ldots \rangle \notin \mathbf{C}$

(Ghidini and Giunchiglia, 2001, p.4)

Satisfiablity then can be defined as follows:

Let $\mathbf{C} = \{\mathbf{c}\}$, with $\mathbf{c} = \langle \mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_i \ldots \rangle$, be a model and $i: \phi$ a formula. $\mathbf{C}$ satisfies $i: \phi$, in symbols $\mathbf{C} \models i: \phi$, if for all $\mathbf{c} \in \mathbf{C}$, $\mathbf{c}_i \models \phi$ where if, for all $m \in$, $m \models_{cl} \phi$.
(Ghidini and Giunchiglia, 2001, p.8)

Furthermore,

A formula $i: \phi$ is valid, *in symbols* $\models i: \phi$, if all models satisfy $i{:}\phi$.

Ghidini and Giunchiglia (2001, p.30f) provide proofs that their system that allows multiple contexts, *Multicontext system*, is complete and sound with respect to a certain model.

5. Buvaç and Mason (1993)

Buvaç and Mason (1993) (cf. Buvaç, Buvaç and Mason 1995) propose a logic that formalises McCarthy's theory of context, which is called Propositional Logic of Context (henceforth PLC). Buvaç and Mason (1993) show that their logic is complete and sound.

Supposing that contexts can be denoted by labels, a set of such labels $\mathbb{K}$ and a set of atomic propositions $\mathbb{P}$, together with the modality *ist* (κ, ϕ) for each $\kappa \in \mathbb{K}$. A set of well-formed formulae $\mathbb{W}$ will be

$$\mathbb{W} := \mathbb{P} \cup (\neg \mathbb{P}) \cup (\mathbb{P} \supset \mathbb{P}) \cup ist\,(\mathbb{K}, \mathbb{P})$$

In order to express a context seen from another context, sequences of contexts are defined as follows. Supposing $\mathbb{K}^*$ denote the set of finite context sequences and $\bar{\kappa} = \kappa_1 \ldots \kappa_n$ denote any element of K*. Then a vocabulary **Vocab** $(\bar{\kappa},\phi)$ can be defined as $\{\langle\bar{\kappa}, p\rangle\}$

A model *M* will then be defined as a relation between a set of partial truth assignments to context sequences:

$M \in (\mathbb{K}^* \rightarrow_p \mathbf{P}\ (\mathbb{P} \rightarrow_p \{\text{true,false}\}))$ where $A \rightarrow_p B$ denotes a set of partial functions from A to B and $\mathbf{P}\ (A)$ denotes the powerset of A.

ϕ is *valid* in a context sequence $\bar{\kappa}$ if $\models_{\bar{\kappa}} \phi$; ϕ is *satisfiable* in a context sequence if there is a PLC-model $\boldsymbol{M}$ such that $\boldsymbol{M} \models_{\bar{\kappa}} \phi$.

Bouquet and Serafini (2000) observe that LMS is more general than PLC: cf. Buvaç and Mason, 1993, for the latter can be embedded in the former. Furthermore, they state that PLC with different vocabularies for different contexts is equivalent to PLC with a single vocabulary for all contexts (p.23). Even if their argument is correct, it does not follow that McCarthy's theory of context in itself is incapable of allowing different vocabularies. And as we have seen in 2, McCarthy's own work makes it clear that the converse is true. There is a far more problematic issue concerning the axiom, Δ, which enables one knowledge base to access another knowledge base and which, as a result, might deny partiality that underlies the theory. As this is more of a logical problem and it is possible to have a propositional logic of context without the axiom, I shall not discuss this further: see de Paiva (2003) for a more detailed discussion.

6. Buvaç (1995)

Buvaç (1995) offers an account of lexical ambiguity which is based on McCarthy's theory of context. He provides a proof theory that has the following properties:

(K) $\vdash \kappa$: ist $(\kappa, \phi \rightarrow \varphi) \rightarrow$ ist (κ', φ)
[Every context is closed w.r.t. logical consequence.]

(Δ) $\vdash \kappa$: ist $(\kappa_1$, ist $(\kappa_2, \phi) \vee \varphi) \rightarrow$ ist $(\kappa_1$, ist $(\kappa_2, \phi)) \vee$ ist (κ_1, φ)
[Contextual omniscience]

(Flat) $\vdash \kappa$: ist $(\kappa_2$, ist $(\kappa_1, \phi)) \rightarrow$ ist $(\kappa_1, \phi))$
[Every context looks the same regardless of which context it is being viewed from.]

(Enter) $$\frac{\vdash \kappa' : \text{ist}\ (\kappa, \phi)}{\vdash \kappa : \phi}$$

(Exit) $$\frac{\vdash \kappa : \phi}{\vdash \kappa' : \text{ist}\ (\kappa, \phi)}$$

The Δ axiom, as mentioned in Section 5, is controversial, and so is Flatness. As the logic without these two axioms is conceivable, this does not count as a real obstacle for using the theory for natural language semantics.

7. Discussion

We have briefly examined three different frameworks. Each of them has potential problems if used for analysing natural language utterances. One reason for this is none of the above mentioned theories incorporates quantification. Another issue is they have rather different logical properties and are not strictly comparable. For instance, Buvaç and Mason (1993) use modal logic. Ghidini and Giunchiglia's theory (2000 and 2001) is centred around the concepts of locality and compatibility whilst McCarthy's original theory is couched in the predicate *ist*, which is basically validity. Furthermore, none of the extant logics of context does not explicate how contexts are obtained; they are simply 'given', which would not be a satisfactory explanation in natural language semantics.

More research is required to decide which framework is most appropriate for analysing and representing natural language expressions. In particular, it would be useful if one could compare computational complexity of each logic, for unnecessary complexity is not desirable even though simplistic logics might not be able to provide sufficient mechanisms.

References

Attardi, G. and M. Simi: 1993, 'A Formalization of Viewpoints', TR-93-062, Universita di Pisa, Pisa.

Bouquet, P. and L. Serafini: 2000, 'Comparing Formal Theories of Context in AI', Technical Report #0201-02, Istituto Trentino di Cultura, Trento.

Buvaç, S.: 1995, 'Resolving Lexical Ambiguity Using a Formal Theory of Context', available from http://www-formal.stanford.edu/buvac.

Buvaç, S., V. Buvaç and I. Mason: 1995, 'Metamathematics of Contexts', *Fundamenta Informaticae* **23**(3), 412-419.

Buvaç, S and I. Mason: 1993, 'Propositional Logic of Context', in R. Fikes and W. Lehnert (eds.), *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pp.412-419, Menlo Park, CA.

Ghidini, C. and F. Giunchiglia: 2001, 'Local Models Semantics, or Contextual Reasoning = Locality + Compatibility', Technical Report, Istituto Trentino di Cultura.

Ghidini, C. and F. Giunchiglia: 2002, 'What is Local Models Semantics?', Technical Report DIT-02-0090, University of Trento.

Guha, R. and J. McCarthy: 2003, 'Varieties of Contexts', paper presented at ConText03, Stanford University, Stanford, CA.

McCarthy, J.: 1983, 'Notes on Formalizing Context', in *IJCAI-93*, pp.555-560, San Mateo, CA.

McCarthy, J.: 1996, 'A Logical AI Approach to Context', MS. Stanford University, Stanford, CA.

McCarthy, J. and S. Buvaç: 1997, 'Formalizing Context (Expanded Notes)', in A. Aliseda *et al.* (eds.), *Computing Natural Language*, pp. 13-50, CSLI Publications, Stanford, CA.

de Paiva, V.: 2003, Natural Deduction and Context as (Constructive) Modality, MS, PARC, Palo Alto, CA.