Quantificational Logic of Context

Saša Buvač
Computer Science Department
Stanford University
Stanford, CA 94305-9020
buvac@cs.stanford.edu
http://www-formal.stanford.edu/buvac

Abstract

In this paper we extend the Propositional Logic of Context, (Buvač & Mason 1993; Buvač, Buvač, & Mason 1995), to the quantificational (predicate calculus) case. This extension is important in the declarative representation of knowledge for two reasons. Firstly, since contexts are objects in the semantics which can be denoted by terms in the language and which can be quantified over, the extension enables us to express arbitrary first-order properties of contexts. Secondly, since the extended language is no longer only propositional, we can express that an arbitrary predicate calculus formula is true in a context. The paper describes the syntax and the semantics of a quantificational language of context, gives a Hilbert style formal system, and outlines a proof of the system's completeness.

Introduction

Contexts first appeared in declarative AI when they were presented as a possible solution to the problem of generality in McCarthy's Turing Award Paper, (McCarthy 1987). Since then, contexts have found uses in various AI applications, including:

- managing large knowledge bases (Guha 1991),
- translating knowledge (Buvač & Fikes 1995),
- modeling knowledge and belief (Giunchiglia 1993),
- integrating data bases (Farquhar *et al.* 1995),
- planning (Buvač & McCarthy 1996),
- qualitative reasoning (Nayak 1994), and
- common sense reasoning (McCarthy & Buvač 1994).

These applications require the expressive power of first-order logics. However, till now no formal logical investigations of quantificational theories of context have been done. The aim of this paper is to rectify this deficiency by extending the Propositional Logic of Context, (Buvač & Mason 1993; Buvač, Buvač, & Mason 1995), to the quantificational (predicate calculus) case. This extension is important in the declarative representation of knowledge for two reasons. Firstly, since contexts are objects in the semantics which can be denoted by terms in the language and which can be quantified over, the extension enables us to express arbitrary first-order properties of contexts. Secondly, since the extended language is no longer only propositional, we can express that an arbitrary predicate calculus formula is true in a context. This paper describes the syntax and the semantics of a quantificational language of context, gives a Hilbert style formal system, and outlines a proof of the system's completeness.

The Logic

We extend classical 2-sorted predicate calculus with identity to enable representing facts about contexts and reasoning with contexts. Our logic has the following four basic features.

1. Contexts are treated as formal objects, i.e. objects in the semantics which can be denoted by terms in the language and which we can quantify over. Consequently, we can state first-order properties of contexts in the same way we state properties of any other objects.
2. The language is extended with a new modality, $\texttt{ist}(k, \phi)$, ($\texttt{ist}$ is pronounced "is true"). It is used to express that the predicate calculus formula, ϕ, is true in the context denoted by the term k.
3. Rather than being given in isolation, all formulas are stated in some context. We write $k : \phi$ when formula ϕ is given in the context denoted by the term k.
4. The formal system contains rules for entering and exiting a context; the proofs which use these rules mirror the intuitive patterns of contextual reasoning.

Semantically, a context is modeled by a set of truth assignments that describe the possible states of affairs of that context. Thus a model will associate a set of first-order structures with every context. These first-order structures reflect the states of affairs which are possible in that context. For an atom to be true in a context, it has to be satisfied by all the structures associated with that context. Therefore, the `ist` modality is interpreted as validity: $\mathtt{ist}(k, \rho)$ is true iff the atom ρ is true in all the first-order structures associated with context k. Treatment of `ist` as validity also corresponds to Guha's proposal for context semantics, which was motivated by the Cyc knowledge base.

The formal system captures some intuitive patterns of contextual reasoning. Intuitively, to prove that a formula is true in some context, we want to first enter that context, perform some inferences with the assumptions made in that context to derive our goal formula, and finally exit the context. When sub-formulas of a formula we want to prove pertain to different contexts, we derive the sub-formulas in their corresponding contexts, and then put them together in the original context to obtain the desired formula. To capture this style of reasoning, we define derivability as a relation on a formula ϕ given in a context k, and write $\vdash k : \phi$. We also introduce the inference rules (**Enter**) and (**Exit**) which enable the reasoning system to enter and exit a context. In (McCarthy 1993), McCarthy illustrates how such rules can be used to generate the desired pattern of reasoning.

We proceed to present some technical details of the logic and a brief sketch the aspects of the completeness proof which are new to the quantificational case. We use standard mathematical notation; we use $\mathbf{P}(X)$ to refer to the set of subsets of X. To simplify the formulas we will distinguish between context objects and non-context objects by assuming two disjoint sorts: the *context sort* and the non-context sort. The latter is referred to as the *discourse sort*, reflecting the intuition that the non-context objects will be the topic of discourse.

Syntax

A language $\mathcal{L}$ of our logic is any language of classical 2-sorted predicate calculus with identity. Formally, language $\mathcal{L}$ is a collection of the constants and the predicates of all arities. We call the sorts the context sort and the discourse sort.

We now fix some language $\mathcal{L}$. The set of all terms in our logic, $\mathbb{T}$, is identical to the set of terms of classical 2-sorted predicate calculus with identity over the same language $\mathcal{L}$. Formally, $\mathbb{T}$ is the set of variables and constants (of both sorts) of the language $\mathcal{L}$. We use $\mathbb{K}$ to refer to the set of terms of the context sort, and $\mathbb{V}$ to refer to the set of variables of both sorts. Note that for simplicity of presentation, our logic has no functions.

The set of atomic formulas, $\mathbb{W}_0$, is the set of atomic formulas of classical 2-sorted predicate calculus with identity: non-logical predicates and the identity predicate applied to an appropriate number of arguments of appropriate sorts. The set $\mathbb{W}$, of well-formed formulas (wffs) of our logic, is defined as the least set satisfying

$$\mathbb{W} := \mathbb{W}_0 \cup (\neg\mathbb{W}) \cup (\mathbb{W} \to \mathbb{W}) \cup (\forall\,\mathbb{V})\mathbb{W} \cup \mathtt{ist}(\mathbb{K}, \mathbb{W})$$

The operations $\wedge$, $\vee$, $\leftrightarrow$, and quantifier $\exists$ are assumed to be defined as abbreviations in the usual way. We will use $\mathbb{W}_{\mathrm{PC}}$ to refer to the set of well-formed formulas of classical predicate calculus with identity, i.e. formulas which do not contain the `ist` modality. To simplify presentation, we assume that the set of bound variables is disjoint from the set of free variables.

We adopt the following notational conventions: $a, a_1, \ldots$ range over constants; $v, v_1, \ldots$ range over variables; $t, t_1, \ldots$ range over terms; $k, k_1, \ldots$ range over terms of only the context sort; and $p, p_1, \ldots$ range over predicates. Lower case Greek letters range over $\mathbb{W}$. The letter T ranges over (possibly infinite, possibly empty) sets of wffs. Note that since all the formulas we will be concerned with are well-formed, the sorts of terms will often be obvious, and will thus not need to be stated explicitly. Similarly, we often do not explicitly list all the arguments of predicates.

Semantics

We begin by fixing some language $\mathcal{L}$, and defining $\mathrm{STR}(\mathcal{L})$ to be the collection of classical 2-sorted first-order structures $<<C, D>, \mathcal{I}>$, i.e. C and D are non empty sets, and $\mathcal{I}$ is standard two-sorted interpretation function for the language $\mathcal{L}$. Intuitively, the set C should be interpreted as the set of context objects, and the set D should be interpreted as the set of discourse objects of the particular structure.

By convention, gothic letters will range over elements of $\mathrm{STR}(\mathcal{L})$. If $\mathfrak{A} = <<C, D>, \mathcal{I}>$, then we use $\mathcal{I}(\mathfrak{A})$ to refer to $\mathcal{I}$, the interpretation function of the first-order structure $\mathfrak{A}$; we use $|\mathfrak{A}|^c$ to refer to C, the set of context objects in the domain of the first-order structure $\mathfrak{A}$; we use $|\mathfrak{A}|^d$ to refer to D, the set of discourse objects in the domain of the first-order structure $\mathfrak{A}$; and we use $|\mathfrak{A}|$ to refer to $C \cup D$, the set of all objects in the domain of the first-order structure $\mathfrak{A}$.

Definition ($\mathfrak{M}$): A model, $\mathfrak{M}$, is a function which maps each context object to a (possibly empty) set of 2-sorted first-order structures of the language $\mathcal{L}$,

$$\mathfrak{M} : \mathrm{Dom}(\mathfrak{M}) \to \mathbf{P}(\mathrm{STR}(\mathcal{L})),$$

provided the following conditions hold:

1. The domains of all first-order structures of all contexts are the same. Formally, for any two context objects c_1 and c_2, for any first-order structures $\mathfrak{A} \in \mathfrak{M}(c_1)$ and $\mathfrak{B} \in \mathfrak{M}(c_2)$, $|\mathfrak{A}|^c = |\mathfrak{B}|^c$ and $|\mathfrak{A}|^d = |\mathfrak{B}|^d$. We use $|\mathfrak{M}|^c$ to refer to the set of context objects $|\mathfrak{A}|^c$, we use $|\mathfrak{M}|^d$ to refer to the set of discourse objects $|\mathfrak{A}|^d$, and we use $|\mathfrak{M}|$ to refer to the set of all objects $|\mathfrak{A}|$. By convention, c,c_1,... range over $|\mathfrak{M}|^c$; d,d_1,... range over $|\mathfrak{M}|^d$; and e,e_1,... range over $|\mathfrak{M}|$.

2. The set of context objects, $|\mathfrak{M}|^c$, is disjoint from the set of discourse objects, $|\mathfrak{M}|^d$.

3. The domain of the model, Dom($\mathfrak{M}$), is identified with the set of context objects, $|\mathfrak{M}|^c$.

4. We require that all interpretation functions map a constant to the same object; we say that all constants are *rigid designators*. Formally, for any $\mathfrak{A} \in \mathfrak{M}(c_1)$ and $\mathfrak{B} \in \mathfrak{M}(c_2)$, and for any constant a, we have $\mathcal{I}(\mathfrak{A})(a) = \mathcal{I}(\mathfrak{B})(a)$.

Definition (variable assignment): A variable assignment is a function from the set of variables, $\mathbb{V}$, to the set of all objects, $|\mathfrak{M}|$, providing variables are assigned objects of appropriate sorts. We extend the variable assignment to constants; this is trivial since all the constants are rigid designators.

By convention, the Greek letters σ and τ will range over variable assignments. Instead of writing $\sigma(v)$, we will use the common notation and write $v[\sigma]$.

We introduce $\models$, which is a relation on a model, a first-order structure, a context, a formula and a variable assignment. The relation $\models$, which is written $\mathfrak{M}, \mathfrak{A} \models k : \phi\,[\sigma]$, should be interpreted as a *satisfaction relation*: we say that the model $\mathfrak{M}$, the first-order structure $\mathfrak{A}$, and the variable assignment σ satisfy the formula ϕ in context $k[\sigma]$.

Definition ($\models$): If $\mathfrak{A} \in \mathfrak{M}(k[\sigma])$ then $\mathfrak{M}, \mathfrak{A} \models k : \chi\,[\sigma]$ is defined by induction on the structure on χ, as follows:

$\mathfrak{M}, \mathfrak{A} \models k : p(t_1, \ldots, t_i)\,[\sigma]$ if $< \mathcal{I}(\mathfrak{A})(t_1[\sigma]), \ldots, \mathcal{I}(\mathfrak{A})(t_i[\sigma]) > \in \mathcal{I}(\mathfrak{A})(p)$

$\mathfrak{M}, \mathfrak{A} \models k : t_1 = t_2\,[\sigma]$ if $\mathcal{I}(\mathfrak{A})(t_1[\sigma]) = \mathcal{I}(\mathfrak{A})(t_2[\sigma])$

$\mathfrak{M}, \mathfrak{A} \models k : \neg\phi\,[\sigma]$ if not $\mathfrak{M}, \mathfrak{A} \models k : \phi\,[\sigma]$

$\mathfrak{M}, \mathfrak{A} \models k : \phi \rightarrow \psi\,[\sigma]$ if $\mathfrak{M}, \mathfrak{A} \models k : \phi\,[\sigma]$ implies $\mathfrak{M}, \mathfrak{A} \models k : \psi\,[\sigma]$

$\mathfrak{M}, \mathfrak{A} \models k : (\forall v)\phi\,[\sigma]$ if for all $e \in |\mathfrak{A}|$ of the same sort as v $\mathfrak{M}, \mathfrak{A} \models k : \phi\,[\sigma(v := e)]$

$\mathfrak{M}, \mathfrak{A} \models k : \mathtt{ist}(k', \phi)\,[\sigma]$ if for all $\mathfrak{B} \in \mathfrak{M}(k'[\sigma])$ $\mathfrak{M}, \mathfrak{B} \models k' : \phi\,[\sigma]$

We write $\models k : \phi$ iff $(\forall \mathfrak{M})(\forall \mathfrak{A} \in \mathfrak{M}(k))\ (\forall \sigma)$ $\mathfrak{M}, \mathfrak{A} \models k : \phi\,[\sigma]$. We call this relation *validity*.

Note that in the clause for universal quantification, the term k can not be the variable v since we have assumed that the set of bound variables is disjoint from the set of free variables.

Formal System

Since all formulas in our logic are given in some context (rather than being given in isolation) derivability is a relation on a formula and a context. We write $\vdash k : \phi$ and say that formula ϕ is *derivable* in context k. We define derivability in a Hilbert style.

Definition (derivability): $\vdash k : \phi$ iff $k : \phi$ is an instance of an axiom schema or follows from provable formulas by one of the inference rules. Formally, $\vdash k : \phi$ iff there is a sequence $[k^1 : \phi^1, \ldots, k^m : \phi^m]$ such that $k^m = k$ and $\phi^m = \phi$ and for each $i \leq m$ either $k^i : \phi^i$ (1) is an instance of one of the axiom schemas, or (2) follows from earlier elements in the sequence via one of the inference rules.

The axiom schemas and inference rules naturally divide into three groups.

1. Classical Predicate Calculus.

(PL) $k : \phi$ provided ϕ is a propositional tautology.

(UI) $k : (\forall v)\phi(v) \rightarrow \phi(t)$

(t =) $k : t = t$

(p =) $k : (t_i = t_i') \rightarrow (p(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) \rightarrow p(t_1, \ldots, t_{i-1}, t_i', t_{i+1}, \ldots, t_n))$

(MP) $\dfrac{k : \phi \quad k : \phi \rightarrow \psi}{k : \psi}$

(UG) $\dfrac{k : \phi \rightarrow \psi(v)}{k : \phi \rightarrow (\forall v')\psi(v')}$ provided v is not free in ϕ.

2. Propositional Properties of Contexts.

(K) $k : \mathtt{ist}(k', \phi \rightarrow \psi) \rightarrow (\mathtt{ist}(k', \phi) \rightarrow \mathtt{ist}(k', \psi))$

(Δ) $k : \mathtt{ist}(k_1, \mathtt{ist}(k_2, \phi)) \vee \mathtt{ist}(k_1, \neg\mathtt{ist}(k_2, \phi))$

(Flat) $k : \mathtt{ist}(k_1, \mathtt{ist}(k_2, \phi)) \leftrightarrow \mathtt{ist}(k_2, \phi)$

(Enter) $\dfrac{k' : \mathtt{ist}(k, \phi)}{k : \phi}$ (Exit) $\dfrac{k : \phi}{k' : \mathtt{ist}(k, \phi)}$

3. Quantificational Properties of Contexts.

(BF) $k : (\forall v)\mathtt{ist}(k', \phi) \rightarrow \mathtt{ist}(k', (\forall v)\phi)$

($\mathtt{ist}$ =) $k : (t_1 = t_2) \leftrightarrow \mathtt{ist}(k', t_1 = t_2)$

We briefly comment on the axioms and inference rules. The first set of axioms and rules guarantees that all valid formulas of classical predicate calculus with identity hold in every context, and every context is closed with respect to the classical rules of inference. The second set of axioms and rules captures the propositional modal properties of contextual reasoning. The axiom schema (**K**) guarantees that every context is closed with respect to logical consequence. A property we call *contextual omniscience* is captured by the ($\triangle$) axiom. Intuitively, every context "knows" what is true in every other context. Thus, although a context need not have complete information about what is true in the world, it will have complete information about other contexts' views of the world. If we interpret contexts as knowledge bases, then contextual omniscience states that every knowledge base "can see" into any other knowledge base. The axiom schema (**Flat**) tells us that every context looks the same regardless of which context it is being viewed from. Rules (**Enter**) and (**Exit**) allow the formal system to respectively enter and exit a context. Note that the (**Enter**) rule is the converse of the (**Exit**) rule. The third set of axioms and rules captures the quantificational properties of contexts. The Barcan formula (**BF**) is needed to make the domains of all the first-order structures in all of the contexts be the same. The ($\mathtt{ist}\ =$) tells us that all terms are treated as rigid designators.

Completeness

In this section we state the completeness of the system and outline the proof. The general structure of the completeness proof and a number of lemmas are similar to those of the propositional system presented in (Buvač, Buvač, & Mason 1995). We will demonstrate the aspects of the proof which are novel to the quantificational case.

Theorem (completeness): $\vdash k : \phi$ iff $\models k : \phi$.

The ($\Rightarrow$) direction is the soundness lemma. It is simple to verify that the axioms are sound and that the rules preserve soundness. We proceed by introducing some concepts which will be needed to outline the ($\Leftarrow$) direction of the completeness proof.

Definition (satisfiability): A set of formulas T is *satisfiable in k* iff there exists a model $\mathfrak{M}$, first-order structure $\mathfrak{A}$, and a variable assignment σ such that for all $\phi \in \mathrm{T}$ $\mathfrak{M}, \mathfrak{A} \models k : \phi[\sigma]$.

Definition (consistency): A formula ϕ is *consistent in k* iff not $\vdash k : \neg\phi$. A finite set of formulas T is consistent in k iff $\bigwedge\!\!\bigwedge \mathrm{T}$, the conjunction of all the formulas in T, is consistent in k. An infinite set T is consistent in k iff every finite subset of T is consistent in k. A set T is inconsistent in k iff the set T is not consistent in k. A set T is maximally consistent in k iff T is consistent in k and for all ϕ if $\phi \notin \mathrm{T}$, then $\mathrm{T} \cup \{\phi\}$ is inconsistent in k.

Definition (ω-completeness): A set of sentences T is *ω-complete* iff for any formula ϕ we have $(\forall v)\phi(v) \notin \mathrm{T}$ implies $\neg\phi(t) \in \mathrm{T}$ for some term t.

Given a set of formulas T in k_0, we will define the set T_k to be those formulas from T which "talk only about" the particular context k, (in the sense that they are true in that context and they contain no $\mathtt{ist}$ modalities). The set of formulas T_k, will be used to define the part of the model of T which describes the state of affairs in the context k.

Definition (T_k): If T is a set of formulas given in k_0, then $\mathrm{T}_k := \{\phi \mid \mathtt{ist}(k, \phi) \in \mathrm{T}$ and $\phi \in \mathbb{W}_{\mathrm{PC}}\}$. We say that T_k is defined from T in k_0.

As is usual, an important part of the completeness proof is the Lindenbaum lemma, allowing any consistent set of wffs to be extended to a maximally consistent set. Our completeness proof will be based on a Henkin construction, which means that in parallel to the process which makes the set maximally consistent, we will also provide witnesses for all previously un-witnessed existential formulas. In a Henkin construction, it is standard to expand the language of the original theory with some infinite set of new constants and use these as witnesses. Our method is similar: previously unused variables will be witnesses. The advantages of using variables for witnesses is that we do not need to change the language of the original set of sentences. We simply need an infinite supply of unused variables.

Notational convention (unused variables): Assume a set of formulas T is given in some context k_0. We use $\tilde{\mathbb{V}}^c$ to denote an infinite set of new context variables, and $\tilde{\mathbb{V}}^d$ an infinite set of new discourse variables. We use $\tilde{v}, \tilde{v}_1, \ldots$ to range over $\tilde{\mathbb{V}}^c \cup \tilde{\mathbb{V}}^d$.

Lemma (Lindenbaum): Assume the set of formulas T_0 is given in k_0, and that the variables in $\tilde{\mathbb{V}}^c$ and $\tilde{\mathbb{V}}^d$ are not used in T_0 or in k_0. If T_0 is consistent in k_0, then T_0 can be extended to a maximally consistent set T in k_0, such that every non-empty set T_k (defined from T in k_0) is ω-complete.

We proceed to outline the proof that any set of formulas T_0 which is consistent in k_0 must be satisfiable in k_0. It is simple to show that this is equivalent to the ($\Leftarrow$) of the completeness theorem.

Proof (completeness): Assume T_0 is consistent in k_0. We extend the set of variables $\mathbb{V}$ used in T_0 and k_0 with an infinite set of new context variables $\tilde{\mathbb{V}}^c$ and an infinite set of new discourse variables $\tilde{\mathbb{V}}^d$ which do not occur in T_0 or in k_0.

By the (**Lindenbaum lemma**) we can extend T_0 to a maximally consistent set T in k_0. Using T we will construct the model $\mathfrak{M}$ from terms of the language $\mathcal{L}$. We identify the sets of objects with the sets of terms: $|\mathfrak{M}|^c$ is defined to be the set of all context terms and $|\mathfrak{M}|^d$ is defined to be the set of all discourse terms. From the maximally consistent set T, for every c, we read off the set T_c. Note that since we identified objects in the model with terms, we are able to interchangeably use terms and objects. Thus we will talk about T_c rather than T_k.

Now we define the set of first order structures associated with a context c. The set of sentences T_c can be thought of as describing the state of affairs which hold at c. We will define the model $\mathfrak{M}$ so that associates the context c with the first-order structures which correspond to this state of affairs. In order to read off these first-order structures we first use T_c to construct $\tilde{\mathrm{T}}_c$, the set of all maximally consistent extensions of T_c. For every set T_c we define $\tilde{\mathrm{T}}_c := \{\mathrm{T} \mid \mathrm{T}$ is a maximally consistent extension of $\mathrm{T}_c\}$. Every $\mathrm{T}' \in \tilde{\mathrm{T}}_c$ will be ω-complete since $\mathrm{T}' \subseteq \mathrm{T}$. Now every $\mathrm{T}' \in \tilde{\mathrm{T}}_c$ is used to read off a first-order structure $\mathfrak{A}$; since T' is maximally consistent and ω-complete this can be done in the usual way. All the first-order structures $\mathfrak{A}$ obtained in this way are put together to define $\mathfrak{M}(c)$, the set which the model $\mathfrak{M}$ will associate to the context c.

This completes the construction of $\mathfrak{M}$. By ($\mathfrak{M}$ **lemma**) we are guaranteed that the model we have constructed is indeed a model.

Lemma ($\mathfrak{M}$): The $\mathfrak{M}$ constructed in the (**completeness proof**) is a model, i.e. it satisfies the additional conditions imposed by the definition of a model.

Finally, to establish completeness we need only show that the model $\mathfrak{M}$ is in fact a model of the sentences T_0 we had started off with. This will be guaranteed by the truth lemma. We define σ_{id} to be the identity function.

Lemma (truth):

$$\mathtt{ist}(c, \phi) \in \mathrm{T} \quad \text{iff} \quad \forall \mathfrak{A} \in \mathfrak{M}(c) \quad \mathfrak{M}, \mathfrak{A} \models c : \phi\,[\sigma_{id}].$$

Clearly, if $\phi \in \mathrm{T}_0$, then also $\phi \in \mathrm{T}$. Since T was given in context k_0, by (**Exit**) rule it follows that $\mathtt{ist}(k_0, \phi) \in \mathrm{T}$, and therefore by the (**truth lemma**) we get $\mathfrak{M}, \mathfrak{A} \models k_0 : \phi\,[\sigma]$. This completes the outline of the completeness proof. $\square_{\textbf{completeness}}$

The proof of the truth lemma is similar to its propositional counterpart. The only new case, that for universal quantifiers, follows simply since having both directions of the Barcan formula enables us to "pull out" all the quantifiers from within an `ist`.

To construct a model the standard Henkin construction needs a number of interesting modifications. We illustrate these by outlining the proof of the Lindenbaum lemma. But first we need to state some simple properties of consistency.

Lemma (consistency): If T is consistent in k_0, then for any wff ϕ

1. at least one of: $\mathrm{T} \cup \{\phi\}$, $\mathrm{T} \cup \{\neg\phi\}$ is consistent in k_0;
2. if $\mathrm{T} \cup \{\neg \mathtt{ist}(k, (\forall v)\phi(v))\}$ is consistent in k_0, then $\mathrm{T} \cup \{\mathtt{ist}(k, \phi(v_1))\}$ is also consistent in k_0, provided v_1 does not occur free in T, k, or ψ.

Proof (Lindenbaum): Our proof of the Lindenbaum lemma is based on a Henkin construction. We enumerate all the sentences in the language $\mathcal{L}$: $\phi_0, \phi_1, \ldots$ and construct an increasing sequence of consistent sets $\mathrm{T}_0 \subset \mathrm{T}_1 \subset \mathrm{T}_2 \subset \cdots$ of sentences of $\mathcal{L}$ such that:

1. Each T_i is consistent.
2. $\phi_i \in \mathrm{T}_{i+1}$ or $\neg\phi_i \in \mathrm{T}_{i+1}$.
3. If $\phi_i = \mathtt{ist}(k, (\forall v)\phi(v))$ and $\neg\phi_i \in \mathrm{T}_{i+1}$, then $\mathtt{ist}(k, \phi(\tilde{v}_p)) \in \mathrm{T}_{i+1}$, where $\tilde{v}_p$ is the first variable from $\tilde{\mathbb{V}}^c$ or $\tilde{\mathbb{V}}^d$ (depending on its sort) not occurring in T_i, ϕ_i, or k.

Now we will construct this sequence of sets of sentences, and prove that it has the above properties 1–3. The construction proceeds in two stages. Assuming we already have the set T_i, we will first construct a temporary set T'_i which will take care of condition 2. Then, in the second stage, using this temporary set T'_i we construct T_{i+1} by adding witness axioms thus also satisfying condition 3.

We elaborate the first stage. Let

$$\mathrm{T}'_i = \begin{cases} \mathrm{T}_i \cup \{\phi_i\} & \text{if } \mathrm{T}_i \cup \{\phi_i\} \text{ is consistent in } k_0 \\ \mathrm{T}_i \cup \{\neg\phi_i\} & \text{if } \mathrm{T}_i \cup \{\neg\phi_i\} \text{ is consistent in } k_0 \end{cases}$$

Note that from the fact that T is consistent in k_0 and the (**consistency lemma 1**) it follows that one of the two choices above has to be consistent. Therefore, this takes care of condition 2. above. Note that often it will be the case that more than one of the above choices is consistent. In this case we can arbitrarily choose which sentence will be added to T_i. However different choices will lead to the construction of different maximally consistent sets which gives us some control over the models we create.

Now we elaborate the second stage of the construction. Let T_{i+1} be $\mathrm{T}'_i \cup \{\mathtt{ist}(k, \phi(\tilde{v}_p))\}$ if $\phi_i = \mathtt{ist}(k, (\forall v)\phi(v))\&\neg\phi_i \in \mathrm{T}'_i$, and T'_i otherwise, where $\tilde{v}_p$ is the first variable from $\tilde{\mathbb{V}}^c$ or $\tilde{\mathbb{V}}^d$ (depending on its sort) not occurring in T_i, ϕ_i, or k. This clearly takes care of the conditions 3. above. All that remains to be shown is that condition 1. holds, i.e. that the set T_{i+1} is consistent. Clearly the set T'_i produced after the first stage of the construction is consistent. By (**consistency lemma 2**) the second stage also produces a consistent set. Therefore, condition 1. holds. Thus we have shown that the sequence of theories which we have constructed has properties 1–3 given above.

We define the set $\mathrm{T} := \cup_{i=0}^{\infty}\mathrm{T}_i$. It is straight forward to show that it is maximally consistent. $\Box_{\textbf{Lindenbaum}}$

A Simple Example

We proceed to illustrate that the quantificational features introduced in this paper are necessary in order to represent real world knowledge in the framework of context logics.

Assume the page for 06/23/96 in McCarthy's diary contains the formula

$$\mathit{fly}(\mathit{UA}, 921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21)$$

which is intended to mean that on 06/23/96 McCarthy is scheduled to fly to Los Angeles on flight 921, leaving San Francisco at 7:00 and arriving in LA at 8:21. Note that although the term *McCarthy* is not mentioned in the above formula, the entry implicitly pertains to McCarthy since the formula is given in McCarthy's diary. Similarly, although the term 06/23/96 is not mentioned, the entry implicitly pertains to the date 06/23/96 since the formula is entered in the diary page associated with that particular date.

One of the original motivations for context formalisms was to aid in expressing such implicit assumptions without having to modify the formula itself (as is proposed in (McCarthy 1993)). Unfortunately, the propositional language of context is not expressive enough to handle even this simple example.

The quantificational language of context is, however, useful in addressing this problem. Firstly, since contexts are objects in the semantics which can be denoted by terms in the language and which can be quantified over, we can express arbitrary first-order properties of contexts. Secondly, since the extended language is no longer only propositional, we can express that an arbitrary predicate calculus formula is true in a context.

These two features in place has allowed the quantificational language of context to describe the implicit assumptions of a formula without modifying the formula itself. The first feature allows us to state properties of the context associated with a dated page in McCarthy's diary: $\lambda c.\mathit{diary}(c, \mathit{McCarthy}) \wedge \mathit{date}(c, 06/23/96)$. The second feature allows us to state that McCarthy's flight information is given in the context of that particular page: $\mathtt{ist}(c, \mathit{fly}(\mathit{UA}, 921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21))$. Putting these together, we get

$$(\forall c)(\mathit{diary}(c, \mathit{McCarthy}) \wedge \mathit{date}(c, 06/23/96)) \rightarrow$$

$$\mathtt{ist}(c, \mathit{fly}(\mathit{UA}, 921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21)).$$

Related Work

This line of research is primarily influenced by McCarthy's notions of context (McCarthy 1987; 1993). The key idea in McCarthy's proposal is to treat contexts as formal objects, which enables one to state properties of contexts and relations on contexts. Also due to McCarthy is the formula $\mathtt{ist}(c, \phi)$, which expresses that formula ϕ is true in context c. The propositional logic of context, (Buvač & Mason 1993; Buvač, Buvač, & Mason 1995), provided the basic formal analysis which this paper extends to the quantificational case.

A comparison of the propositional logic of context to other formalizations of context in AI and to multimodal logics is given in (Buvač, Buvač, & Mason 1995). The key points of the comparison to the formalizations in AI (Lifschitz 1986; Guha 1991; Shoham 1991; Giunchiglia 1993; Nayak 1994; Attardi & Simi 1995) carry over to the quantificational logic of context. However, the comparison of the propositional logic of context to propositional multi-modal logics does not carry over to the quantificational case. Thus we proceed to compare the quantificational logic of context to multi-modal logics.

Comparison to Multi-Modal Logics

There is a clear parallel between the logic of context and the standard multi-modal logics, like the ones used for reasoning about knowledge and belief of multiple agents (Halpern & Moses 1992). In the propositional case, given a context language containing a set of contexts $\mathbb{K}$, we can define a modal language containing modalities $\Box_1, \Box_2, \ldots$, one for each context from $\mathbb{K} = \{k_\beta\}_{\beta<\alpha}$. By replacing each occurrence of $\mathtt{ist}(k_\beta, \psi)$ with $\Box_\beta\psi$, we can define a bijective translation function which to each formula of the propositional context logic assigns a well-formed modal formula. Based on this translation, (Buvač, Buvač, &

Mason 1995) shows a reduction of the propositional logic of context to a propositional multi-modal logic.

However, these results do not carry over to the quantificational case. The quantificational logic of context, for example, enables us to state that the formula ψ is true in contexts which satisfy some property $p(x)$ as follows:

$$(\forall v)p(v) \rightarrow \mathtt{ist}(v, \psi).$$

This formula has no obvious translation into any standard multi-modal logic. The meaning of such formulas which quantify over modalities is beyond the analysis commonly done in quantificational modal logic.

Our derivability relation, $\vdash k : \phi$, differs from the usual modal logic derivability relation, $\vdash \phi$. This choice was influenced by the intuition that every formula is given in some context and that the reasoning system can enter and exit a context. If we were willing to give up these features, we could define derivability in the style that is standard to modal logics.

Conclusion

Our main motivation for formalizing contexts is to solve the problem of generality in AI. We want to be able to make 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. Such a capability would allow the designer of a reasoning system to include only such phenomena as are required for the system's immediate purpose, while retaining the assurance that if a broader system is required later, "lifting axioms" can be devised to restate the facts from the narrower context to the broader one, with qualifications added as necessary. Thus, a necessary step in the direction of addressing the problem of generality in AI is providing a language which enables representing and reasoning with multiple contexts and expressing lifting axioms. In this paper we provide such a language.

Acknowledgments

The author thanks Ian A. Mason for his help throughout this research. The author would also like to thank Richard Fikes, R. V. Guha, John McCarthy, Jeaneah Paik, and Johan van Benthem for their valuable comments. This research is supported in part by the Advanced Research Projects Agency, ARPA Order 8607, monitored by NASA Ames Research Center under grant NAG 2-581, by NASA Ames Research Center under grant NCC 2-537, Navy N66001-94-D-6052, and Darpa contract NAG2-703.

References

Attardi, G., and Simi, M. 1995. A formalization of viewpoints. *Fundamenta Informaticae* 23(3).

Buvač, S., and Fikes, R. 1995. A declarative formalization of knowledge translation. In *Proceedings of the ACM CIKM: The 4th International Conference on Information and Knowledge Management.*

Buvač, S., and Mason, I. A. 1993. Propositional logic of context. In *Proceedings of the Eleventh National Conference on Artificial Intelligence.*

Buvač, S., and McCarthy, J. 1996. Combining planning contexts. In Tate, A., ed., *Advanced Planning Technology–Technological Achievements of the ARPA/Rome Laboratory Planning Initiative.* AAAI Press.

Buvač, S.; Buvač, V.; and Mason, I. A. 1995. Metamathematics of contexts. *Fundamenta Informaticae* 23(3).

Farquhar, A.; Dappert, A.; Fikes, R.; and Pratt, W. 1995. Integrating information sources using context logic. Technical Report KSL-95-12, Stanford.

Giunchiglia, F. 1993. Contextual reasoning. *Epistemologia, special issue on I Linguaggi e le Macchine* XVI.

Guha, R. V. 1991. *Contexts: A Formalization and Some Applications.* Ph.D. Dissertation, Stanford.

Halpern, J. Y., and Moses, Y. 1992. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence* 54.

Lifschitz, V. 1986. On formalizing contexts. Unpublished manuscript.

McCarthy, J., and Buvač, S. 1994. Formalizing Context (Expanded Notes). Technical Note STAN-CS-TN-94-13, Stanford.

McCarthy, J. 1987. Generality in artificial intelligence. *CACM* 30(12).

McCarthy, J. 1993. Notes on formalizing context. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence.*

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

Shoham, Y. 1991. Varieties of context. In Lifschitz, V., ed., *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy.* Academic Press.