MODELING CONTEXT WITH SITUATIONS

Mehmet Surav and Varol Akman

BILKENT UNIVERSITY

Department of Computer Engineering and Information Science

Technical Report BU-CEIS-95-07

Modeling Context with Situations

Mehmet Surav and **Varol Akman**
Department of Computer Engineering and Information Science
Bilkent University, Bilkent, Ankara 06533, Turkey
Phone: +90 (312) 266 4133 (sec.)
Fax: +90 (312) 266 4126
E-mail: {surav,akman}@bilkent.edu.tr

Abstract

The issue of context arises in assorted areas of Artificial Intelligence. Although its importance is realized by various researchers, there is not much work towards a useful formalization. In this paper, we will present a preliminary model (based on Situation Theory) and give examples to show the use of context in various fields, and the advantages gained by the acceptance of our proposal.

Keywords: Context, Commonsense Reasoning, Situation Theory.

For the time being, all I want you to grant is that context is important—so important that we have to represent it clearly in the knowledge base. Dana Scott [14, p. 354]

1 Introduction

Although the term *context* is frequently used in explanations, proofs, etc. in Artificial Intelligence, its meaning is left to the reader's understanding, i.e., it is used in an implicit and intuitive manner [1]. However, when we are to implement a system, we have to make this notion explicit using, hopefully, a formal approach.

In this work, our aim is to offer a useful formalization of context, one that can be used for automated reasoning in Artificial Intelligence, Computational Linguistics, and so on. To this end, we will first review logic-based attempts towards formalizing context. In general, the focus of our discussion will be around McCarthy's proposal [12].

Our approach, on the other hand, is inspired by a pioneering work of Barwise [3] and will be presented using the notation and terminology of Situation Theory. We will give the necessary background to Situation Theory, and review the contributions of Barwise. Then we will advance our proposal and discuss the handles that it offers on the issue of context. We will present examples, mostly taken from the available literature, so that we convince the reader that our formalization is quite useful.

2 Previous Formalizations in Logic

The notion of context was first introduced to AI in a logicist framework by McCarthy in his 1971 Turing Award talk. (This talk was later published as [11].) After that introduction, research on the topic was quite silent until the late eighties. McCarthy published his recent ideas on context in [12]. Other notable works on formalizing context are due to Guha [10], Shoham [15], Giunchiglia [9], S. Buvač and Mason [6], and Attardi and Simi [2].

In his most recent work [12], McCarthy states three reasons for introducing the *formal* notion of context.

- The use of context allows simple axiomatizations. He exemplifies this by stating that axioms for blocks world situations can be *lifted* to contexts involving fewer assumptions.
- Contexts allow us to use a specific vocabulary and information about a circumstance.
- 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.

The basic relation relating contexts and propositions is $\mathtt{ist}(c, p)$. It asserts that proposition p is true in context c. Then the main formulas are sentences of the form

$$c' : \quad \mathtt{ist}(c, p) \tag{1}$$

In other words, p is true in context c, and this itself is asserted in an outer context c'.

Some properties of context, which will form a base for our formalization, include:

1. *Contexts are abstract objects.*
   McCarthy says [12, p. 1]: "We do not offer a definition [of context], but we will offer some examples." Some contexts will be *rich* objects, e.g., the situations in Situation Calculus. Some contexts will not be as rich, e.g., some simple micro-theories [10].

2. *Contexts are first-class objects.*
   We can use contexts in our formulas in the same way we use other objects.

3. *There are some relations working between contexts.*
   The most notable one is the *more general than* ($\preceq$) relation. This defines a partial ordering over contexts. Using $\preceq$, we can lift a fact from a context to one of its super-contexts using the following nonmonotonic rule:
   $$\forall c_1 \, \forall c_2 \, \forall p \;\; (c_1 \preceq c_2) \wedge \mathtt{ist}(c_1, p) \wedge \neg ab1(c_1, c_2, p) \rightarrow \mathtt{ist}(c_2, p)$$
   Here, c_2 is a super-context of c_1 and p is a predicate of c_1, $ab1$ is an abnormality predicate and $\neg ab1(c_1, c_2, p)$ is used to support the nonmonotonicity. In other words, the above rule is a basic lifting rule from a context to its super-context. (Obviously, we can state a similar rule between a context and one of its sub-contexts.)

4. *There are some functions to form new contexts by specialization.*
   One example McCarthy uses is the function *specialize-time*(t, c) which returns a context related to context c in which time is specialized to have the value t.

5. *Lifting rules.*
   According to McCarthy [12], the main goal of the use of contexts is to simplify axiomatizations (by allowing us to lift axioms from one context to another). Lifting rules are always asserted in an outer context which should be capable of supporting such rules.

Using lifting rules, we can do the following while we are transferring an axiom:

1. *No operation.*
   If two contexts are using the same terminology for a concept, this is a natural choice. For example, the following lifting rule states that we can use the axioms related to $on(x, y)$ relation of *above-theory* context in *general-blocks-world* context without any change:
   $$c_0 : \forall x \forall y \; \mathtt{ist}(\textit{above-theory}, on(x, y)) \rightarrow \mathtt{ist}(\textit{general-blocks-world}, on(x, y))$$

2. *Change the arity of a predicate.*
   In different contexts, the same predicate might take a different number of arguments. McCarthy's example for this is *on* which takes two arguments in *above-theory* context, and three arguments in a context c in which *on* has a third argument denoting the situation[1]. The lifting rule is
   $$c_0 : \forall x \forall y \forall s \; \mathtt{ist}(\textit{above-theory}, on(x, y)) \rightarrow \mathtt{ist}(\textit{context-of(s)}, on(x, y, s))$$
   where *context-of* returns the context associated with situation s in which the usual *above-theory* axioms hold.

3. *Change the name of a predicate.*
   Similar to the case with arities, we can change the name of a predicate via lifting rules. For example, we can translate *on* to *üzerinde*, when we move from *above-theory* to *turkish-above-theory* [2]:
   $$c_0 : \forall x \forall y \; \mathtt{ist}(\textit{above-theory}, on(x, y)) \rightarrow \mathtt{ist}(\textit{turkish-above-theory}, \textit{üzerinde}(x, y))$$

[1] Here the word "situation" is used in the Situation Calculus sense.

[2] Note that, in the above examples, the lifting rules are always stated in an outer context, c_0, so that `ist` formulas can be used without any paradoxical (circular) side effects. Attardi and Simi [2] criticize McCarthy for his unclear use of lifting rules, and prove that if a condition for stating lifting rules in outer contexts is not asserted, lifting rules might introduce paradoxes.

When we take contexts in the natural deduction sense (as per McCarthy's suggestion [11]), the operations of *entering* and *leaving* a context might be useful and shorten the proofs involving contexts. In this case, $\mathtt{ist}(c, p)$ will be analogous to $c \rightarrow p$, and the operation of entering c can be taken as $assuming(p, c)$. Then, entering c and inferring p will be equivalent to $\mathtt{ist}(c, p)$ in the outer context.

Important achievements on the formalization of context include the works of Attardi and Simi [2], and Giunchiglia [9]; these use natural deduction as the reasoning machinery. Attardi and Simi [2] offer a "viewpoint" representation in which contexts are considered to be sets of reified sentences of the FOL. In Giunchiglia [9, 5], the notion of *MultiContext (MC) System* is introduced. An MC system is defined as a pair $< \{c_i\}_{i \in I}, \Delta >$, where $\{c_i\}_{i \in I}$ is the set of contexts and Δ is the set of *bridge rules*. Here, context is a triple $c_i =< L_i, A_i, \Delta_i >$ where L_i is the language of c_i, A_i is the set of axioms of c_i, and Δ_i is the set of inference rules that can be used only in c_i. Bridge rules are the inference rules, similar to the lifting rules, linking different contexts. They are of the form

$$\frac{< A, C_1 >}{< B, C_2 >} \quad ,$$

allowing us to derive a formula B in context C_2 from the formula A in context C_1.

Relative decontextualization is another issue raised by McCarthy's work. He proposes a mechanism of relative decontextualization to do the work of eternal sentences. The mechanism depends on the premise that when several concepts 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 the Problem Solving Contexts (PSC) of CYC [10].

Another place where context might be useful is a *mental state* [12]. McCarthy thinks of mental states as outer contexts. The advantage of representing mental states as outer contexts is that we can include the reasons for having a *belief*. Then, when we are required to do *belief revision*, the inclusion of the reasons for having a belief simplifies our work. When we use beliefs as usual (i.e., no belief revision is required), we simply enter the related context and use them.

3 Situation Theory (Barwise on Contexts)

Situation Theory is a principled programme to develop a unified mathematical theory of meaning and information content, and to apply that theory to specific areas of language, computation, and cognition. Barwise and Perry [4] claim that for an expression to have meaning, it should convey information. They develop a theory of situations and of meaning as a relation between situations. The theory provides a system of abstract objects that make it possible to describe the meaning of both expressions and mental states in terms of the information they carry about the external world [7].

The two major concepts of Situation Theory are infons and situations. Infons are the basic informational units and are denoted as $\ll P, a_1, \ldots a_n, i \gg$ where P is an n-place relation, $a_1, \ldots a_n$ are objects appropriate for the respective argument places of P, and i is the polarity (0 or 1).

Situations are first-class citizens of the theory, and are defined intensionally. A situation is considered to be a structured part of the reality that an agent (somehow) manages to pick out.

It is desirable to have some computational tools to handle situations. Abstract situations are the mathematical constructs which are amenable to mathematical manipulation. An abstract situation is defined as a (possibly non-well-founded) set of infons. Given a real situation s, the set $\{\alpha \mid s \models \alpha\}$ is the corresponding abstract situation. Here, s *supports* α (denoted as $s \models \alpha$) means that α is an infon that is true of s.

Related to parametric infons, there is a construct by which we can assign "values" to parameters. Formally, an *anchor* for a set, A, of basic parameters is a function defined on A, which assigns to each parameter T_i in A an object of type T. For example, if f anchors $\dot{a}$ to the individual "Sullivan," we write $f(\dot{a}) =$ Sullivan to denote this anchoring.

Object-types are determined over some initial situation. Let s be a given situation. If $\dot{x}$ is a parameter and I is a set of infons (involving $\dot{x}$), then there is a type $[\dot{x}|s \models I]$. This is the type of all those objects to which $\dot{x}$ may be anchored in s, for which the conditions imposed by

I obtain. We refer to this process of obtaining a type from a parameter $\dot{x}$, a situation s, and a set I of infons, as *type-abstraction.* $\dot{x}$ is known as the *abstraction parameter* and s is known as the *grounding situation.*

In Situation Theory, the flow of information is realized via *constraints*. We represent a constraint as $\ll involves, S_0, S_1, 1 \gg$ where S_0 and S_1 are situation-types between which the information is carried out. Cognitively, if this relation holds, then it is a fact that if S_0 is realized (i.e., there is a real situation $s_0 : S_0$), then so is S_1 (i.e., there is a real situation $s_1 : S_1$). For example,we represent the regularity "Smoke means fire" with the following constraint c:

$$S_0 = [\dot{s} | \dot{s} \models \ll \textit{smoke-present}, \dot{l}, \dot{t}, 1 \gg]$$
$$S_1 = [\dot{s} | \dot{s} \models \ll \textit{fire-present}, \dot{l}, \dot{t}, 1 \gg]$$
$$c = \ll involves, S_0, S_1, 1 \gg$$

In Situation Theory information is taken into account in an atomic fashion, i.e., information at the level of the relation provided by the individuation schema. In other words, information is regarded as essentially *propositional*: information is information *about* some situation s. This either has the form $s : S$ where s is a situation and S is a situation-type, or else $x : T$ where x is an object and T is an object-type defined over some grounding situation u. In the latter case, if $T = [\dot{x} | u \models \sigma]$ then $x : T$ if and only if $u \models \sigma[f]$ where $f(\dot{x}) = x$.

Barwise's ideas on circumstance, thus on context, arise from his work on conditionals and circumstantial information [3]. Barwise states two innovative examples, one on a missing pollen and the other on the (wrong) proof of $1 = -1$. We will only mention the missing pollen example.

Example: The Missing Pollen

Let us consider Claire (Barwise's then nine-month old daughter). Barwise knows that if Claire rubs her eyes, then she is sleepy. This is expressed by the conditional

If Claire rubs her eyes, then she is sleepy.

For months, this was a sound piece of (conditional) knowledge that Barwise and his wife used to understand Claire, and learn when they should put her to bed. However, in the early summer, it began to fail them. Combined with other symptoms, they eventually figured out that Claire was allergic to something or other. They called it "Pollen X" since they did not know its precise identity. So Pollen X could also cause Claire to rub her eyes.

Barwise formally approaches the above problem as follows. Briefly, with constraint $C = [S \Rightarrow S']$, a real situation s contains information relative to such an actual constraint C, if $s : S$. Clearly, s may contain various pieces of information relative to C, but the most general proposition that s contains, relative to C, is that s' is realized, where $s' : S'$.

Thus, Barwise represents the information that "If Claire rubs her eyes, then she is sleepy" with the following parametric constraint C:

$$S = [\dot{s} | \dot{s} \models \ll rubs, \text{Claire, Claire's eyes}, \dot{l}, \dot{t}, 1 \gg]$$
$$S' = [\dot{s} | \dot{s} \models \ll sleepy, \text{Claire}, \dot{l}, \dot{t}, 1 \gg]$$
$$C = \ll involves, S, S', 1 \gg$$

Before Pollen X was present, the above constraint represented a reasonable account. However, when Pollen X arrived, the constraint became inadequate and required revision. Barwise points out to two alternatives to this end:

- From [if ϕ then ψ] infer [if ϕ and β, then ψ].
- From [if ϕ then ψ] infer [if β, then if ϕ then ψ].

Here β corresponds to the additional background conditions.

Barwise chooses the second way to deal with the problem[3], modifies *involves*, and makes the background assumptions explicit by introducing a third parameter. The relation now becomes:

$$\ll involves, S_1, S_2, B, 1 \gg$$

[3]Although these alternatives are equivalent from a logical point of view, the second is more appropriate to reflect the intuitions behind the background conditions. In the first case, the rule *if ϕ then ψ* is directly modified to use background conditions whereas, in the second case, it is evaluated only when the background conditions hold.

For example, with the new *involves*, the Missing Pollen Example can be solved via the introduction of a background condition B, which supports the following:

$$\ll exists, \text{Pollen X}, \dot{l}, \dot{t}, 0 \gg$$

We can thus reformulate the problem using B:

$$S = [\dot{s} | \dot{s} \models \ll rubs, \text{Claire, Claire's eyes}, \dot{l}, \dot{t}, 1 \gg]$$
$$S' = [\dot{s} | \dot{s} \models \ll sleepy, \text{Claire}, \dot{l}, \dot{t}, 1 \gg]$$
$$B = [\dot{s} | \dot{s} \models \ll exists, \text{Pollen X}, \dot{l}, \dot{t}, 0 \gg]$$
$$C = \ll involves, S, S', B, 1 \gg$$

4 A Formalization of Context in Situation Theory

The main purposes of our proposal may be categorized into three:

1. To offer a representation schema which allows contexts in a uniform way.
2. To support the essential properties of context.
3. To clarify the notion of context and to reduce it to a mathematical problem. (As a future project, the computer implementation of the proposal might be considered [16].)

We will approach context as an amalgation of grounding situation and the rules which govern the relations within the context. Thus we will represent a context by a situation type which supports two types of infons:

- *factual* infons to state the facts and the usual bindings. This infons might be saturated or unsaturated depending on the completeness of that piece of knowledge.
- *conditional* infons (which correspond to parametric conditionals) representing infons to capture the *if-then* relations and axioms within the context.

Thus, the context of an M.S. Thesis Presentation can be formulated with the following situation-theoretic constructs. Let c be Sullivan's M.S. Thesis Presentation context. This context supports some infons to represent the basic facts, and constitutes the basis for parameter binding. Some of the infons are

$$c \models \ll school, \text{Bilkent}, 1 \gg \tag{2}$$
$$c \models \ll department, \text{Computer Science}, 1 \gg \tag{3}$$
$$c \models \ll ms\text{-}student, \text{Sullivan}, 1 \gg \tag{4}$$
$$c \models \ll ms\text{-}advisor, \text{Ackerman}, 1 \gg \tag{5}$$
$$c \models \ll ms\text{-}jury\text{-}member, \text{Necker}, 1 \gg \tag{6}$$

Within this context, we have some natural regularities valid for all thesis presentation contexts, such as c_1:

$$S_1 \models \ll ms\text{-}advisor, \dot{a}, 1 \gg \tag{7}$$
$$S_2 \models \ll ms\text{-}jury\text{-}member, \dot{a}, 1 \gg \tag{8}$$
$$c_1 = \quad [S_1 \Rightarrow S_2 | B] \tag{9}$$

Here B is a background situation such that

$$B \models \ll school, \text{Bilkent}, 1 \gg \tag{10}$$

The constraint c_1 can be represented with the following infon:

$$c_1 \models \ll involves, S_1, S_2, B \gg \tag{11}$$

The second part, i.e., the set of infons in Equations 7–11, intuitively states that thesis advisors of are also jury members in M.S. Thesis Presentation contexts (in Bilkent). Using the above context as a grounding situation with the anchoring $f(\dot{a}) = \text{Ackerman}$, we can conclude that Ackerman is also a jury member.

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

A crucial property is that contexts are first-class objects, so that we can use them in the same way as other objects. In our approach, we are modeling contexts with situation types, and situation types are situations which have some unbound parameters. Other than having unbound parameters, situation types are ordinary situations, and thus first class objects of Situation Theory. (Having unbound parameters does not cause any problem.)

Richness of the contexts was stated by McCarthy [11, 12] and Guha [10]. In Situation Theory, situations are, by definition, rich objects [7]. The richness of situations leads to the partiality of contexts as McCarthy notes.

Another aspect of the use of context is the flexibility of having private rules and presuppositions related to a particular viewpoint. In the logicist approach, presuppositions were represented with predicates which contained no variables (either bound or unbound) and rules were represented with quantified logical implications:

$$c: \quad present(Air) \tag{12}$$

$$c: \quad \forall x\, bird(x) \rightarrow flies(x) \tag{13}$$

Equation 12 states that air is present (a presupposition), and Equation 13 states that if something is a bird, it flies (a default rule). The same capability is also available in our notion of context. We represent the facts related to a particular context with parameter free infons supported by the situation type which corresponds to the context. The rules of the context are represented by the constraints. Since constraints are allowed to be parametric, we can easily use them as rules related to the context. The above example might then be restated as follows:

$$S_c \models \ll present, \mathrm{Air}, 1 \gg \tag{14}$$

$$\begin{array}{l} S_1 = [\dot{s} | \dot{s} \models \ll bird, \dot{a}, 1 \gg \\ S_2 = [\dot{s} | \dot{s} \models \ll flies, \dot{a}, 1 \gg \\ S_c \models \ll involves, S_1, S_2, 1 \gg \end{array} \tag{15}$$

Here, the fact that air is present is represented with the infon in Equation 14, and the rule that birds fly is represented with the constraint in Equation 15. Therefore, we can use the situation type S_c as as the context of the logicist approach (namely, context c).

A related issue is the background information. Barwise [3] points out to the importance of background information in the *involves* relation. In our model, the context representation is designed to supply the adequate background information.

Consider the context of an M.S. Thesis presentation. In this context, the advisor of the thesis is also a jury member of the thesis. This rule is stated via c_1 in Equation 9. In the constraint, the background condition is that the school is Bilkent University. This is stated by using B of Equation 10. In our definition of context, supplying this kind of background information is simple (and in fact necessary).

Contexts define the domain of quantification. This property of context is due to its use as a grounding situation, so that in the binding of parameters, the only available objects are those available in the context.

By lifting (or using bridge rules), we can use some axioms from one context in another context. Lifting rules (whether nonmonotonic or not) are always stated in the outer one of the two contexts between which the lifting will be done. In our case, lifting has similar properties. Basically, we will state lifting rules as constraints. Nonmonotonicity of the lifting will be realized by the background conditions in the *involves* relation.

Let C_1 and C_2 be the contexts between which the lifting is to be done. Let C be the outer context. Let us state a lifting rule (C below) to lift relation *foo* from context C_1 to relation *bar* in context C_2:

$$S_1 = \ [\dot{s} |\ C_1 \models \ll foo, \dot{a}, 1 \gg] \tag{16}$$

$$S_2 = \ [\dot{s} |\ C_2 \models \ll bar, \dot{a}, 1 \gg] \tag{17}$$

$$C \models \ll involves, S_1, S_2, B, 1 \gg \tag{18}$$

(B is the background condition, which enables us to have nonmonotonicity while lifting.)

Regarding lifting, there are some discrepancies between our approach and the logicist one. Namely, in the logicist approach we can change the arity of a relation while lifting; in our approach this is not allowed. This is not due to a limitation on our part, but is rather a by-product of the philosophy behind Situation Theory. In Situation Theory an individuation mechanism is used to name objects, individuals, events, situations, and so on. One application area of the individuation mechanism is relations. For example, once we individuate the *on* relation with two parameters as

$$\ll on, \dot{a}, \dot{b}, 1 \gg \tag{19}$$

we always consider *on* with two parameters, i.e., in all situations and groundings we use *on* with this fixed number of parameters. In the logicist way, *on* is taken in a syntactic sense, and in some contexts it might require two parameters, while in some other contexts it might require three. (For example, the third parameter might correspond to time.) Although Situation Theory seems to be weaker at first regard, it in fact gives us the mechanisms to compensate for this weakness. In the *on* example, we can compensate the requirement for time in one context by simply stating an infon, which enables us to represent the dependence on time.

McCarthy's Lifting Example

In [12], McCarthy states the following example in a subsection titled "Lifting Rules." Here, we will first present the original example, and then re-do it in our version of formal context.

The Original Example

McCarthy considers two contexts, namely, *Above-Theory (AT)* and c. *Above-Theory* contains some simple blocks-world assumptions, similar to Equations 20 and 21. In AT, the notion of situation is undefined. However, c supports the situations, and the predicates usually have an additional parameter for the situation. For example $on(x, y)$ becomes $on(x, y, s)$, where s corresponds to the situation in which $on(x, y)$ holds. In c, *context-of*(s) is a function, which returns a specialization (a sub-context) of c, where the situation is fixed to s. The lifting rules working between c and one of its specializations are written as Equations 22 and 23. Equation 24 is the major lifting axiom, which links AT and the sub-contexts of c. The example in McCarthy [12] is that from $\mathtt{ist}(c, on(A, B, S_0))$ we can prove $\mathtt{ist}(c, above(A, B, S_0))$. In the proof, c_0 is the outer context. The axioms are the following:

$$c_0 :\ AT :\ \forall x\, \forall y\ \ on(x, y) \rightarrow above(x, y) \tag{20}$$

$$c_0 :\ AT :\ \forall x\, \forall y\, \forall z\ \ above(x, y) \wedge above(y, z) \rightarrow above(x, z) \tag{21}$$

$$c_0 :\ c :\ \forall x\, \forall y\, \forall s\ \ on(x, y, s) \leftrightarrow \mathtt{ist}(context\text{-}of(s), on(x, y)) \tag{22}$$

$$c_0 :\ c :\ \forall x\, \forall y\, \forall s\ \ above(x, y, s) \leftrightarrow \mathtt{ist}(context\text{-}of(s), above(x, y)) \tag{23}$$

$$c_0 :\ c :\ \forall p\, \forall s\ \ \mathtt{ist}(AT, p) \rightarrow \mathtt{ist}(context\text{-}of(s), p) \tag{24}$$

The proof proceeds as follows:

$$c_0 :\ c :\ on(A, B, S_0) \tag{25}$$

$$c_0 :\ c :\ \mathtt{ist}(context\text{-}of(S_0), on(A, B)) \tag{26}$$

$$c_0 :\ c :\ context\text{-}of(S_0) :\ on(A, B) \tag{27}$$

$$c_0 :\ c :\ \mathtt{ist}(context\text{-}of(S_0), \forall x\, \forall y\ on(x, y) \rightarrow above(x, y)) \tag{28}$$

$$c_0 :\ c :\ context\text{-}of(S_0) :\ above(A, B) \tag{29}$$

$$c_0 :\ c :\ above(A, B, S_0) \tag{30}$$

Equation 25 is the assumption given in the beginning. Equation 26 is obtained from Equations 22 and 25 by binding A to x, B to y, and S_0 to s. Equation 27 obtained from Equation 26 by entering the context *context-of*(S_0). Equation 28 is the result of lifting Equation 20 by the lifting axiom (Equation 24). From Equations 27 and 28, we obtain Equation 29. The desired result (Equation 30 is obtained from Equations 29 and 23. This proof is summarized in Figure 1 where contexts are represented as Venn diagrams. Atomic formulas are represented with capital letters, and transfers between contexts (i.e., the lifting rules) are represented by arrows. We have labeled arrows in the way the proof grows. Basically, McCarthy is drawing a virtual link from the atomic formula X to the atomic formula V. Since, c has no rule to draw

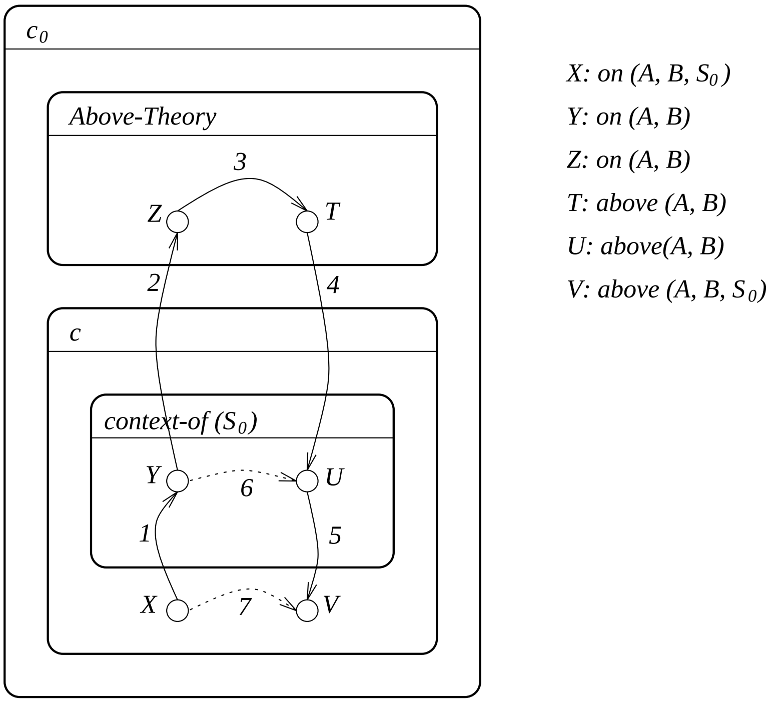

Figure 1: Diagram of McCarthy's proof

a link from X to V, we first create the context *context-of* (S_0), and draw a link to the atomic formula Y using the lifting rule in Equation 22. After Y, McCarthy lifts the implication of $above(x, y)$ from $on(x, y)$ (the arc labeled with 3 in the figure) to *context-of* (S_0) (i.e., he forms the link 6). Then from Y, by tracing link 6, we get U. From U, by leaving *context-of* (S_0), we get the desired formula V.

As the reader may notice, in the proof of McCarthy, it is more natural to use the path 1-2-3-4-5. However, this path requires one more lifting rule to transfer Y to Z (link 2). In Attardi and Simi [2], this link is explicitly stated and a proof is carried out with the path 1-2-3-4-5.

The Original Example Revisited

The main difference between the statement of McCarthy and ours will be in the language used. We will use Situation Theory to state the example, and will be required to make the following changes:

- Logical implications (such as Equation 20) will be represented with constraints.
- In McCarthy's original example, *on* has different arities in different contexts, i.e., in AT its arity is two whereas in c its arity is three. However, in Situation Theory, we are required to refer to *on* in different contexts with different names[4].
- *context-of* (S_0) will be represented with infons of type
$$\ll \textit{context-of}, A, B, 1 \gg$$
where A is a parameter which is of type situation (of Situation Calculus) and B is a parameter of type situation (of Situation Theory), which corresponds to A.
- The contexts c_0, c, and AT of McCarthy will be represented with the contexts S_{c0}, c_c, and c_{AT}, respectively. In fact, all of the equations in the sequel are supported by the situation S_{c0}. (This situation will be the outermost grounding situation in our proof.)

[4] Once again, the reason behind this change is the schema of individuation used when we identify the *on* relation. Basically, *on* in context AT has no notion of situation, but *on* in context c has this notion. In this case, we cannot refer to the first *on* in the same way as we refer to the second. Thus, we will name our *on* relations as on_{AT} and on_c. Note that we cannot simply regard on_{AT} as on_c with a missing parameter. In Giunchiglia [9], each context has its own language, and therefore *on* may have different semantics in different contexts. However, Situation Theory is a semantical theory, and since *on* has different senses in these different contexts, we have to distinguish them.

- The background conditions $B_{AT}, B_{c\text{-}AT}, B_{AT\text{-}c}$[5] will not be explicitly stated in the proof, for the original proof does not involve any nonmonotonic inference.

The axioms and assumptions of McCarthy will be represented with the following situation-theoretic constructs:

$$\begin{array}{l} S_{11} = [\dot{s} | \dot{s} \models \ll on_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ S_{12} = [\dot{s} | \dot{s} \models \ll above_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ c_{AT} \models \ll involves, S_{11}, S_{12}, B_{AT} \gg \end{array} \tag{31}$$

$$\begin{array}{l} S_{21} = [\dot{s} | \dot{s} \models \ll above_{AT}, \dot{x}, \dot{y}, 1 \gg \;\; \wedge \;\; \dot{s} \models \ll above_{AT}, \dot{y}, \dot{z}, 1 \gg] \\ S_{22} = [\dot{s} | \dot{s} \models \ll above_{AT}, \dot{x}, \dot{z}, 1 \gg] \\ c_{AT} \models \ll involves, S_{21}, S_{22}, B_{AT} \gg \end{array} \tag{32}$$

$$\begin{array}{l} S_{31} = [\dot{s} | \dot{s} \models \ll on_{c}, \dot{x}, \dot{y}, \dot{s}_{31}, 1 \gg] \\ c_{AT} = [\dot{s} | \dot{s} \models \ll on_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ c_{c} \models \ll context\text{-}of, \dot{s}, \dot{s}_{31}, 1 \gg \\ c_{c} \models \ll involves, S_{31}, c_{AT}, B_{c\text{-}AT} \gg \end{array} \tag{33}$$

$$\begin{array}{l} S_{31} = [\dot{s} | \dot{s} \models \ll above_{c}, \dot{x}, \dot{y}, \dot{s}_{41}, 1 \gg] \\ c_{AT} = [\dot{s} | \dot{s} \models \ll above_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ c_{c} \models \ll context\text{-}of, \dot{s}, \dot{s}_{41}, 1 \gg \\ c_{c} \models \ll involves, S_{41}, c_{AT}, B_{c\text{-}AT} \gg \end{array} \tag{34}$$

$$\begin{array}{l} c_{AT} \models \sigma_p \\ c_{c} \models \ll context\text{-}of, \dot{s}, \dot{s}_{51}, 1 \gg \\ \dot{s}_{51} \models \sigma_p \\ c_{c} \models \ll involves, c_{AT}, \dot{s}_{51}, B_{c\text{-}AT} \gg \end{array} \tag{35}$$

In addition to McCarthy's axioms, we will need a further rule to lift facts from c_{AT} to c_c:

$$\begin{array}{l} S_{61} = [\dot{s} | \dot{s} \models \ll above_{c}, \dot{x}, \dot{y}, \dot{s}_{61}, 1 \gg] \\ c_{AT} = [\dot{s} | \dot{s} \models \ll above_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ c_{c} \models \ll context\text{-}of, \dot{s}, \dot{s}_{61}, 1 \gg \\ c_{c} \models \ll involves, c_{AT}, S_{61}, B_{AT\text{-}c} \gg \end{array} \tag{36}$$

In McCarthy, we have

$$c_0 : \mathtt{ist}(c, on(A, B, S_0)). \tag{37}$$

This can be represented with

$$\sigma_0 = \ll on_c, A, B, S_0, 1 \gg$$

and then Equation 37 corresponds to

$$c_c \models \sigma_0. \tag{38}$$

In Situation Theory, this kind of logical proof corresponds to finding the anchoring function, by which we can show that c_c also supports the predicate to be proven, i.e., we must show that

$$c_c \models \ll above_c, A, B, S_0, 1 \gg \tag{39}$$

In the proof, we will first transfer the fact to c_{AT}, then reason that *on* implies *above*, and carry this new fact to c_c. This is the path 1-2-3-4-5 in Figure 1.

Using the constraint in Equation 33 with the anchoring

$$\begin{array}{c} f_1(\dot{s}) = S_0 \\ f_1(\dot{x}) = A \\ f_1(\dot{y}) = B \end{array}$$

[5] B_{AT} is the background condition used in the constraint in AT, $B_{c\text{-}AT}$ is the background condition used in lifting from c to AT, and $B_{AT\text{-}c}$ is the background condition used in lifting from AT to c.

we transfer $\ll on_c, A, B, S_0, 1 \gg$ from c_c to $\ll on_{AT}, A, B, 1 \gg$ in c_{AT}. This corresponds to tracing links 1 and 2 in Figure 1. Note that, we did not lose the parameter S_0, since we will use the same anchoring function when we return to c_c. In c_{AT}, using the anchoring

$$f_2(\dot{x}) = A$$
$$f_2(\dot{y}) = B$$

and Equation 31, we get $\ll above_{AT}, A, B, 1 \gg$. This corresponds to link 3 in Figure 1. After this implication of *above* from *on*, we should transfer the fact to c_c. This is done using Equation 36 with f_1. The result is $\ll above_c, A, B, S_0, 1 \gg$. This completes the proof path 1-2-3-4-5 in Figure 1. Once more, by using one constraint, we have traced two links, namely, 4 and 5, in Figure 1. During this pass, we are in fact referring to U when we use f_1.

Consequently, using two anchoring functions (f_1 grounded at the outermost context S_{c0} and f_2 grounded at c_{AT}), we have carried out the proof of McCarthy in our situation-theoretic framework.

Barwise's Missing Pollen Example Revisited
The following are the constituents of the constraint C, which was the solution to the missing pollen problem:

$$S = [\dot{s} | \dot{s} \models \ll rubs, \text{Claire, Claire's eyes}, \dot{l}, \dot{t}, 1 \gg] \tag{40}$$
$$S' = [\dot{s} | \dot{s} \models \ll sleepy, \text{Claire}, \dot{l}, \dot{t}, 1 \gg] \tag{41}$$
$$B = [\dot{s} | \dot{s} \models \ll exists, \text{Pollen X}, \dot{l}, \dot{t}, 0 \gg] \tag{42}$$
$$C = \ll involves, S, S', B, 1 \gg \tag{43}$$

Initially, it was winter and there were no pollens. The context, call it c_1, must be a situation type which supports

$$c_1 \models \ll exists, \text{Pollen X}, \dot{l}, \dot{t}, 0 \gg$$

(and possibly other things related to Claire, rubbing one's eyes, etc.). Using context c_1 as the grounding situation, we do not violate the background condition B (Equation 42) of constraint C (Equation 43), and thus can conclude that "Claire is sleepy."

Later, in summer, the new context, c_2, supports the infon

$$c_2 \models \ll exists, \text{Pollen X}, \dot{l}, \dot{t}, 1 \gg$$

and when we use c_2 as the grounding situation, we are faced with an inconsistency between B and c_2. Therefore, C becomes void for the new context of the talk, and the conclusion "Claire is sleepy" cannot be reached.

A Nonmonotonicity Example
In Equation 13, if x is somehow bound to a non-flying bird like a penguin, this implication is invalidated using some nonmonotonic technique.

As we have stated before, in Situation Theory, we represent implications with constraints. While stating the constraints, we can use background conditions to add nonmonotonicity:

$$S_1 = [\dot{s} | \ \dot{s} \models \ll bird, \dot{x}, 1 \gg] \tag{44}$$
$$S_2 = [\dot{s} | \ \dot{s} \models \ll flies, \dot{x}, 1 \gg] \tag{45}$$
$$B = [\dot{s} | \dot{s} \models \ll penguin, \dot{x}, 0 \gg \ \wedge \ \dot{s} \models \ll present, \text{Air}, 1 \gg] \tag{46}$$
$$C = \ll involves, S_1, S_2, B, 1 \gg \tag{47}$$

C states that every bird flies unless it is a penguin or there is no air. Here, the important contribution of the situation-theoretic account is that the environmental factors can be easily included in the reasoning phase by suitably varying B.

5 Conclusion

In the literature, there are number of attempts towards a formalization of context in a logicist framework. Our approach differs from these in being stated in the framework of Situation Theory and is primarily an extension of Barwise's conception of context. In [3], Barwise uses

	Mc87	Mc93	Gu91	Ba86	Sh91	Gi93	BM93	AS94	Ours
Logic vs. Situation Theory	Logic	Logic	Logic	S.T.	Logic	Logic	Logic	Logic	S.T.
Modal Treatment	No	No	No	No	Yes	No	No	No	No
Natural Deduction	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No
Paradox-Free	No	No	No	?	?	Yes	Yes	Yes	?

Legend:

Mc87	McCarthy: Generality in AI [11]
Mc93	McCarthy: Notes on formalizing context [12]
Gu91	Guha: A formalization of contexts [10]
Ba86	Barwise: Constraints and conditional information [3]
Sh91	Shoham: Varieties of context [15]
Gi93	Giunchiglia: Contextual Reasoning [9]
BM93	S. Buvač and Mason: Propositional logic of context [6]
AS94	Attardi and Simi: A formalization of viewpoints [2]
Ours	This paper

Table 1: Comparison of the previous approaches and our approach

grounding situations similar to our contexts. However, in his work, content of the grounding situations is not fully described. In our work, we are explicitly stating what a context includes: parameter free infons to state the facts and the usual bindings, and parametric infons to state if-then rules. Table 1 summarizes the essential elements of previous research on context and our proposal.

Compared to other approaches, ours has the following notable properties:

- We might easily require the content of a context change dynamically. We can add (delete) assumptions and rules into (from) a context. Having a dynamic notion of context is not a novel thing for the logicist, since one can always add (delete) axioms into (from) a theory. However, when we fortify our context with dynamic constraints whose background conditions are also dynamic, we get nonmonotonicity in the framework of Situation Theory.

- Since contexts are first-class objects, we can use them in the same way we use the other objects (in the framework of Situation Theory) provided that we are given an appropriate outer context to do this. In this view, our contexts have exactly the same properties as the logicists' contexts.

- Our contexts can also support uncertainty to some extent: we can have unsaturated infons and leave the uncertain part of a piece of information unsaturated so that if we have further knowledge about that piece of information, we can fill the missing portions of an infon.

Obviously, our proposal is not complete. A tentative list for the future work includes:

- *Extension to temporal domain:* In the statement of our proposal, we have not dealt with temporal relations. In Guha's work [10], most of the examples are related to time. As a future project, the study of the temporal relations and information within our contexts might be useful.

- *The need for a Situation Theory tool:* Since we are using a situation theoretic framework, we should have a programming environment for Situation Theory. There are two serious attempts to do this: BABY-SIT [16] and PROSIT [13]. As these attempts progress, the computational aspects of our approach can be better investigated.

- *Some classical problems:* We have not discussed the application of our approach to epistemic puzzles (à la Smullyan). However, a similar study by Ersan and Akman [8] can be adopted to our formalization of context.

Acknowledgments

We are grateful to Fausto Giunchiglia for his expert advice, help, and encouragement. Akman's research is supported in part by a NATO Science for Stability grant (TU-LANGUAGE).

References

[1] V. Akman and E. Tın. What is in a context? In E. Arıkan, editor, *Proceedings of the 1990 Bilkent International Conference on New Trends in Communication, Control, and Signal Processing*, volume II, pages 1670–1676. Elsevier, Amsterdam, Holland, 1990.

[2] G. Attardi and M. Simi. Proofs in context. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourth Conference*, 1994.

[3] J. Barwise. Conditionals and conditional information. In E. C. Traugott, C. A. Ferguson, and J. S. Reilly, editors, *On Conditionals*, pages 21–54. Cambridge University Press, Cambridge, UK, 1986.

[4] J. Barwise and J. Perry. *Situations and Attitudes.* MIT Press, Cambridge, MA, 1983.

[5] P. Bouquet and F. Giunchiglia. Reasoning about theory adequacy: A new solution to the qualification problem. Technical Report 9406-13, Istituto Per La Ricerca Scientifica e Tecnologica, Trento, Italy, 1994. To appear in *Fundamenta Informaticae.*

[6] S. Buvač and I. A. Mason. Propositional logic of context. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pages 412–419, Washington, DC, 1993.

[7] K. Devlin. *Logic and Information.* Cambridge University Press, New York, NY, 1991.

[8] M. Ersan and V. Akman. Situated modeling of epistemic puzzles. *Bulletin of the IGPL*, 1995. To appear.

[9] F. Giunchiglia. Contextual reasoning. *Epistemologia*, XVI:345–364, 1993. Short version in *Proceedings of the IJCAI '93 Workshop on Using Knowledge in Its Context*, Chambery, France, 1993.

[10] R. V. Guha. *Contexts: A Formalization and Some Applications.* PhD thesis, Computer Science Department, Stanford University, Stanford, CA, 1991.

[11] J. McCarthy. Generality in artificial intelligence. *Communications of the ACM*, 30:1030–1035, 1987.

[12] J. McCarthy. Notes on formalizing context. In *Proceedings of the Thirteenth International Joint Conference in Artificial Intelligence (IJCAI '93)*, Chambery, France, 1993.

[13] H. Nakashima, H. Suzuki, P. K. Halvorsen, and S. Peters. Towards a computational interpretation of situation theory. In *Proceedings of the International Conference on Fifth Generation Computer Systems*, pages 489–498, Institute for New Generation Computer Technology, Tokyo, Japan, 1988.

[14] D. S. Scott. Capturing concepts with data structures. In R. A. Meersman and A. C. Sernadas, editors, *Data and Knowledge*, pages 341–354. Elsevier, Amsterdam, Holland, 1988.

[15] Y. Shoham. Varieties of context. In V. Lifschitz, editor, *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy*, pages 393–408. Academic Press, Boston, MA, 1991.

[16] E. Tın and V. Akman. BABY-SIT: A computational medium based on situations. In C. Martín-Vide, editor, *Current Issues in Mathematical Linguistics*, pages 299–308. Elsevier, Amsterdam, Holland, 1994.