CONTEXTS AND SITUATIONS

Mehmet Surav and Varol Akman

BILKENT UNIVERSITY

Department of Computer Engineering
and
Information Science

Technical Report BU-CEIS-94-33

CONTEXTS AND SITUATIONS

Mehmet Surav and Varol Akman
Department of Computer Engineering and Information Science
Bilkent University

October 6, 1994

Abstract

The issue of context arises in assorted areas of Artificial Intelligence, Mathematical Logic, and Natural Language Semantics. Although its importance is realized by various researchers, there is not much work towards a useful formalization. In this report, we will try to identify the problem, and decide what we need for an acceptable (formal) account of the notion of context. 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.

(This is a revised version of the first author's M.S. thesis.)

Keywords: Context, Knowledge Representation, Commonsense Reasoning, Situation Theory and Situation Semantics.

Contents

Table of Symbols

Symbol	Meaning	First Occurrence
$\mathtt{holds}(C,p)$	predicate p is true in context C	p. 7
$\mathtt{ist}(c,p)$	predicate p is true in context c	p. 11
$ab(c_1,c_2,p)$	predicate p is abnormal to lifting from context c_1 to context c_2	p. 11
$\mathit{context\text{-}of}(p)$	returns the context related to p	p. 12
$value(c,t)$	returns the value of term t in context c	p. 12
$\mathit{specialize\text{-}time}(t,c)$	returns the sub-context of context c at time t	p. 13
$\mathit{specializes\text{-}time}(t,c_1,c_2)$	context c_2 is the specialization of context c_1 at time t	p. 13
$assuming(p,c)$	returns a context similar to c in which predicate p is assumed	p. 13
$\mathsf{in}('A',vp)$	sentence A is provable from viewpoint vp	p. 18
$\mathsf{Bel}(g,A)$	agent g has the belief A	p. 19
$\mathsf{True}(A)$	sentence A is true	p. 19
$\mathsf{K}(g,A)$	agent g has the knowledge A	p. 19
$\mathsf{Holds}(A,s)$	sentence A holds in situation s	p. 20
$\preceq$	*more general than* relation	p. 8
p^c	predicate p is true in context c	p. 17

Symbol	Meaning	First Occurrence
$\forall$	for all	p. 11
$\rightarrow$	logical implication	p. 11
$\leftrightarrow$	logical equivalence	p. 12
$\wedge$	logical conjunction	p. 11
$\neg$	logical negation	p. 11
$\dot{\supset}$	more general than	p. 17
$x \dot{\wedge} y$	greatest lower bound of contexts x and y	p. 18
$x \dot{\vee} y$	least upper bound of contexts x and y	p. 18
$\dot{\neg} x$	context not compatible with x	p. 18
$\vdash_C$	derivable without using reflection rules	p. 19
$\ll \cdots \gg$	infon	p. 21
$s \models \sigma$	situation s supports infon σ	p. 22
$[- \mid -]$	situation definition construct	p. 23
$\dot{a} \uparrow \ll ..., \dot{a}, ... \gg$	$\dot{a}$ is a restricted parameter of $\ll ..., \dot{a}, ... \gg$	p. 22
$s : S$	situation s is of type S	p. 23
$\ll involves, S_1, S_2, 1 \gg$	situation type S_1 involves situation type S_2	p. 23
$S_1 \Rightarrow S_2$	situation type S_1 involves situation type S_2	p. 24
$S_1 \Rightarrow S_2 \vert B$	situation type S_1 involves situation type S_2 with background conditions B	p. 25

Chapter 1

Introduction

The issue of context arises in various areas of Artificial Intelligence, Mathematical Logic, and Natural Language Semantics. Although the term *context* is frequently used in explanations, proofs, etc. in these domains, its meaning is left to the reader's understanding, i.e., it is used in an implicit and intuitive manner. However, when we are to implement a system, we have to make this notion explicit using, hopefully, a formal approach.

According to the *Oxford Advanced Learner's Dictionary of Current English* [31, p. 184] the word "context" usually has two meanings: (i) the words around a word, phrase, etc. often used for helping to explain the meaning of the word, phrase, etc. (ii) the general conditions in which an event, action, etc. takes place. Clearly, the first meaning is closely related to linguistic meaning and linguists' use of the word, whereas the second (more general) meaning is the one which is closer to our account of context in this report. In *The Dictionary of Philosophy* [4, p. 47], the word "context" is defined as follows:

> **context** (L. *contexere*, "to weave together." from *con*, "with," and *texere*, "to weave"): The sum total of meanings (associations, ideas, assumptions, preconceptions, etc.) that (a) are intimately related to a thing, (b) provide the origins for, and (c) influence our attitudes, perspectives, judgments, and knowledge of that thing.

From the above definitions, one may form a rough picture of the notion of context. In another dictionary (*Collins Cobuild English Language Dictionary* [19, p. 305]), the associated meanings of "context" include the following:

1. The *context* of something consists of the ideas, situations, events, or information that relate to it and make it possible to understand it fully.

2. If something is seen *in context* or if it is put *into context*, it is considered with all the factors that are related to it rather than just being considered on its own, so that it can be properly understood.

3. If a remark, statement, etc. is taken or quoted *out of context*, it is only considered on its own and the circumstances in which it was said are ignored. It, therefore, seems to mean something different from the meaning that was intended.

As noted by McCarthy [38], context plays an important role in commonsense reasoning. Whenever we state an axiom, we intend to use it in a certain context. If we want to be *general*, then we have to axiomatize in a high level of generality. This results in longer and complicated axioms which can nonetheless be stated in a compact way in a *particular* situation [38, 40]. So, by modeling contexts, we gain two important advantages [3]: (i) *Economical:* We can shorten our axioms, and (ii) *Philosophical:* We can eliminate the difficulties associated with being fully general in the expression of facts, e.g., when we are working on a case which might never occur in real life, we can avoid the effects of this case on our other axioms.

In this report, 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 first identify the role of context in various fields such as Artificial Intelligence, Mathematical Logic, and Natural Language Semantics. Chapter 2 does this and may be considered as an account of the motivation for our study. In Chapter 3, we review some logic-based attempts towards formalizing context. In that chapter, the focus of our discussion will be McCarthy's proposal [40], which, in our view, is the groundwork for all other logicist formalizations[1].

Our approach is inspired by the pioneering works of Barwise [8, 9] and Seligman [49], and will be presented using the notation and terminology of Situation Theory. In Chapter 4, we will give the necessary background to Situation Theory, and review the contributions of Barwise. In Chapter 5, we will advance our proposal, and discuss the handles that it offers on the issue of context. Then we will present examples, mostly taken from the available literature, so that we convince the reader that our formalization is at least as useful as the ones outlined previous approaches.

Finally, in Chapter 6, we will evaluate our approach and discuss its deficiencies. Suggestions for improvement and plans for further research will also be made in this chapter.

[1]In addition to McCarthy, we will review his coworkers' contributions. Attardi and Simi's natural deduction based approach [5, 6] will also be studied in that chapter.

Chapter 2

The Role of Context

In this chapter, we will examine the role of context in various fields, discuss possible applications, and in general try to answer the question "Why should we try to model context?".

Although we focus our attention primarily on Natural Language and Reasoning, we will also review diverse areas some of which are related to Natural Language and Reasoning, some are not. This chapter may appear to be organized in a somewhat haphazard way; we will link the ideas in the upcoming chapters.

2.1 Context in Natural Language

Every natural language utterance occurs in a particular context. The meaning of the utterance and its interpretation, i.e., deciding the content of utterance, its truth value, the information carried out to the addressee by the utterance, and so on, are all context dependent. (Obviously, the 'degree' of context-dependence may vary.)

In this section, we present a simple (possibly trivial for human beings) segment of conversation, and begin to examine the role of context. The skeleton of the example is taken from Barwise [9, p. 27].

A (a woman, talking to B): *I am a philosopher.*

B (talking to C): *She is a philosopher.*

C (talking to A): *So, you are a philosopher.*

In this example, one of the very first context dependent things is the word "philosopher." The meaning of this word is determined using the context of conversation. Although the above excerpt is not sufficient to carry the proper connotation of this word, our common understanding selects an appropriate meaning from a set of possible meanings[1].

In the above example, *indexicals*—"I," "she," and "you"—can be bound to appropriate persons only by the help of context. For example, the sentences uttered by A and B have the same

[1]According to the *Merriam-Webster Dictionary* [56, p. 522], the word "philosopher" may stand for any of the following: a reflexive thinker; a student of or specialist in philosophy; one whose philosophical perspective enables him to meet trouble calmly. In our example, it is the second meaning which is commonly evoked in the minds of the hearers of the word.

content[2], and we can only say this using some circumstantial information and conventions about the conversation[3]. (*Demonstratives*—"that," "this," and so on—have a similar dependency on the circumstance.) This circumstantial information might be formalized via context, so that in reasoning we can propose a formal procedure to deal with it[4].

Another role of context arises when we deal with *quantifiers* [23]. The *range* and *interpretation* of quantifiers depend on the context. For example, the quantifier "all" usually does not apply to all objects, but only to those of a particular kind in a particular domain, determined by the contextual factors. Another example might be the interpretation of the meaning of "many." In an automobile factory, 10 automobiles might not qualify as "many," but if a person owns 10 automobiles it counts as "many" (even the last interpretation is context dependent)[5].

Context might be used to fill the missing parameters of some actions in natural language utterances. Consider an utterance of the sentence

Carl Lewis is running.

In this sentence, the place and time of the action are determined by the context. For example, if we are watching a competing Lewis on TV at the 1992 Barcelona Olympic Games, the place and time of the utterance are different from what we would get if we are watching him practice from our window. Thus, in the first case the place of the "running" action is filled with the Olympic stadium at Barcelona and time of the "running" action is filled with August 1992. In the second case, the place is say, our front lawn and the time is say, June 1992.

In natural language there might be more than one meaning of a word (common examples are "pen" and "bank") and context is the most influential factor in determining the appropriate meaning. In the above utterance, if one does not know who Carl Lewis is, then "running," too, could be interpreted differently.

Some of the natural language relations directly depend upon the context. A good example for this is an utterance of the sentence

Engineering Building is to the left of the Library.

In the general context of Bilkent Campus (cf. Figure 2.1), if we are watching the buildings from the Publishing Company, the utterance is true, but if we are in the Tourism School, the utterance is false. More interestingly, if we are looking from the Rector's Residence, this utterance might be considered neither true nor false[6]. (The Library is behind the Engineering Building.)

Finally, the effects of the environment, that is, the information coming through our five senses must be taken into account as a contextual factor or as a contributor of the context.

[2]The content of an utterance is considered in an intensional manner; namely we suppose the content of all the utterances in the example is the same: "A is a philosopher."

[3]*Anaphora* might be considered as a superset of indexicals. The resolution of anaphora is a complex task because it requires finding the correct antecedent among many possibilities. It involves syntactic, semantic, and pragmatic issues [50, 30], and introduction of the notion of context might give us a more uniform way of dealing with this problem. (Approaches to resolving anaphora using Situation Theory can be found in Gawron and Peters [23], and Tın and Akman [60].)

[4]The representational aspects and the properties of context will be discussed in the following chapters, and after formulating a clear proposal for context, we will re-visit this example in Section 5.2.5.

[5]One might propose that "many" can only be interpreted as a ratio. But even this idea has a contextual dependency on the ratio. In a class of students, half of the students cannot be considered as "many" to cancel a midterm exam, but surely must be regarded as "many" in an influenza epidemic.

[6]We can generalize to *perspectives* from this example [48, 14, 49]. In the following section, we will examine examples similar to this one from a "natural regularity" point of view. This example will also be re-visited after the elaboration of our proposal (cf. Section 5.2.3).

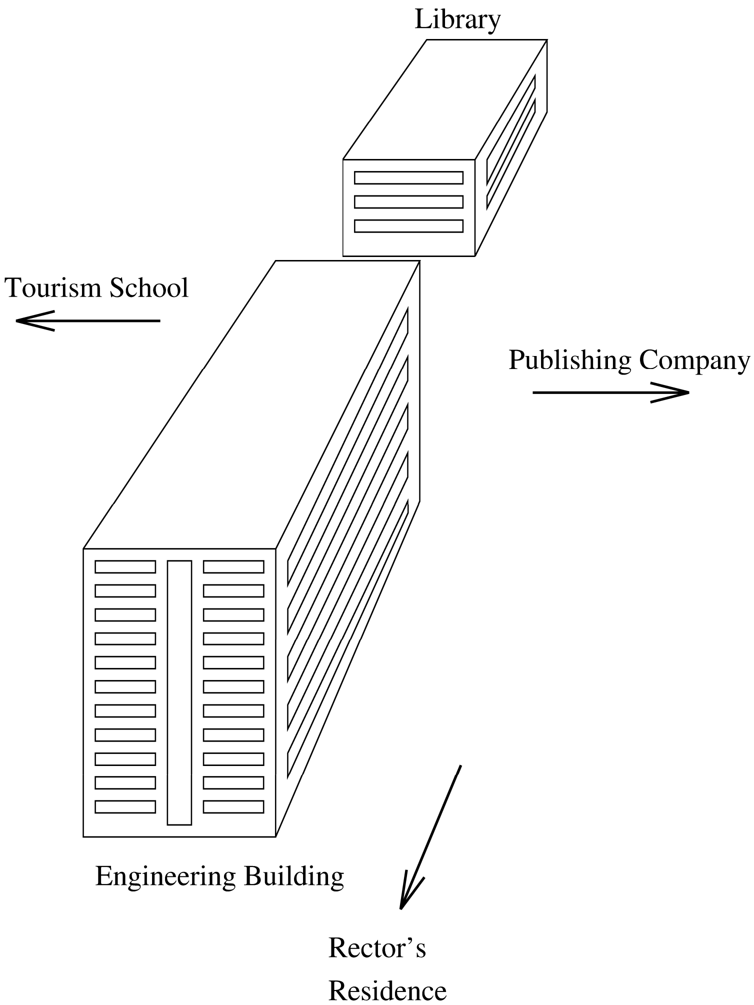

Figure 2.1: A partial view of the Bilkent Campus

After the above identifications, we can conclude that for natural languages a *fleshing-out* strategy [44, 33, 54], i.e., converting everything into decontextualized eternal sentences [43], cannot be employed since we do not always have full and precise information about the relevant circumstances.

2.2 Context in Categorization

Categorization is one of the basic mental processes of reasoning [16, 47]. We, as human beings, can categorize various types of objects, events, situations, etc, and our categorizations depend on the circumstance and perspective. In this section, we will give examples to show the role of context in categorization, and emphasize its connection to natural regularities [14, 49].

In [2], the following example is given to motivate a commonsense set theory:

> In Springfield (home-town of Bart Simpson), there are three barbers working for money, and one barber who does not work for money (since he has another job) but serves the community by shaving senior citizens on Sundays. If we look at the situation from a commonsense perspective, there are four barbers in town, but from say, the mayor's point of view, there are only three (licensed, tax-paying, etc.) barbers.

From the example, it is clear that context (or perspective) plays an important role in classification. In [2] and [55], Akman and I focused our attention on the membership relation (*in*) and introduced a context-dependent membership. However, upon carefully reviewing the literature on natural regularities [48, 14, 11, 49] and categorization [47], we realized that we should begin to deal with this problem not in set theory but in a more philosophical and psychological framework. With the help of [49] and [14], we can transfer the discussion to the domains of Situation Theory and cognitive science. Although our proposal for formalizing context does not offer an alternative way of capturing context in categorization, it uses some of the ideas from previous works [14, 49, 47].

Barwise and Seligman [14] use the analysis of natural regularities to point out to the role of context in categorization. Since categorization is one of the basic processes when we are to deal with regularities, this is a fruitful path for us to analyze the role of context. An example regularity from Seligman [49] is "Swans are white." This is a typical example of natural regularity in the sense that it is both reliable and fallible. Natural regularities are reliable since they are needed to explain successful representation, knowledge, truth, and correct reference. And they are fallible since they are needed to account for misinterpretation, error, false statements, and defeasible reference. Swans are in general white, thus the regularity is reliable and explains a fact. There might be exceptions like the Australian Swans[7], but this does not mean that the regularity does not hold. Here, the fundamental problem with isolating the essential properties of a regularity is that any statement of them depends on some context of evaluation, i.e., we should evaluate the regularity "Swans are white," for say, European swans.

At this point, the reader might notice a correlation between non-monotonic reasoning and the role of context dependent factors in natural regularities [14, 49]. Although natural regularities are usually considered in philosophical discussions, they intuitively correspond to material implication in logic, and the effect of contextual factors is similar to the effect of non-monotonicity. The difference between the philosophical and the logical approaches is in the way that these two disciplines differ. In logic, implication and non-monotonicity are usually studied in a syntactic fashion, and the reasons behind the abnormalities are usually left out of the scope of discussion and/or ignored. On the other hand, in the works of Seligman and Barwise, natural regularities are analyzed from

[7] Australian swans are usually black.

different perspectives, ranging from a purely mathematical approach [11] to a strictly philosophical one [49, 14].

If we could completely describe the *contextual factors,* then the problem would go away and we would not require extra machinery. However, we should always include a "so forth" part to cover the unexpected contextual factors [51]. In many cases, it is simply impossible to state all of the related contextual factors [57, 58, 51]. Still, we must somehow be able to deal with contextual factors. This is why the introduction of some notion of context, and using this notion in categorization might be useful[8].

2.3 Context in Reasoning and AI

In our discussion of the role of context in reasoning and AI, we will mainly study the motivations of McCarthy [40] and his co-workers [24, 52, 18, 5].

When we state something we do so in a context. Similar to our discussion on the role of context in natural language, in reasoning we interpret some predicates differently in different contexts. For example, 37 degrees centigrade is "high" in the context of a weather report, but "normal" in the context of medical diagnosis. In the context of Newtonian mechanics $f = ma$, but in the context of general relativity, this is hardly the case. The examples can be continued. The main point here is that if we are to reason in a common sense way, we have to use certain contexts.

The importance of the notion of context has realized by philosophers for centuries. Early on, philosophers recognized that a casual connection between two events holds only relative to a certain background, thus only in certain contexts. McCarthy was the first researcher to realize that the introduction of a *formal* notion of context is required for "generality" in AI [38].

According to McCarthy [38, 40], there is simply no most general context in which all the stated axioms always hold and everything is meaningful. When we write an axiom, it holds in a certain context, and one can always present another (more general) context in which the axiom does not hold.

Without really introducing a formal notion of context, we may choose to state our axioms in fairly high generality to cover a large domain. But in this case, the axioms become longer. The implementation of CYC, a large commonsense knowledge base and reasoning system, exemplifies this issue [26, 25]. In this case, the system is quite general, and the axioms to cover the real word can only be stated in a lengthy way. The main question is "To what generality should we extend the axioms of the system?". If we commit ourselves to the view that there is no most general context, our system will always have to be partial. At this point, McCarthy proposes the formalization of notion of context. By adding a context parameter to each function and relation, and employing non-monotonic reasoning methods, one can state axioms in a fairly simple way and use them by lifting to other contexts. (Lifting will be explained in Chapter 3.)

Another notion which is discussed by McCarthy is the formalization of relativized-truth-within-a-context via a special predicate. The predicate

$$\texttt{holds}(C, p)$$

is used to state that predicate p holds in context C [9].

[8] Note that, the discussions in [14] and [49] continue with the identification of the problem from the categorization point of view. However, we will not go into the details of that.

[9] In McCarthy's newer work [40], holds is changed to $\texttt{ist}(c, p)$. We will review McCarthy's more recent thinking in the next chapter.

If we compare the two approaches, namely, `holds` and adding a context parameter to each function and predicate, we might choose `holds`, since it allows us to use the context uniformly as the other objects. The major problem with this approach is that, if we are to live in a first order world, we have to somehow reify [38, 24] p in $\texttt{holds}(C, p)$.

Between two contexts, we might consider a *more general than* relation ($C_1 \preceq C_2$) meaning that the second context contains all the information of the first context and probably more. (Intuitively, the second context is broader than the first one.) An example for two contexts related by $\preceq$ is the context of Surav's M.S. Thesis presentation and the context of Bilkent University M.S. Thesis presentations. Here, the second context is more general than the first one; in fact, the first one is obtained from the second by fixing the speaker to Surav, place to A-502, date to September 23rd, etc.

McCarthy proposes that providing operations such as *entering* and *leaving* might be useful for a logical system which uses contexts in a natural deduction sense [61]. However, McCarthy also states that taking contexts as a set of axioms (even as an infinite set of assumptions) is probably incorrect [38].

Among the advantages gained as a result of the use of contexts are the following [52]:

- *Economy of representation:* Different contexts can limit the parts of the knowledge base that are accessible in different ways, effectively allowing the representation of knowledge base in a single structure.
- *Economy of reasoning:* By factoring out a possibly very large knowledge base, context may permit much more efficient reasoning about the real, intended scope.
- *Allowing inconsistent knowledge bases:* The knowledge base might be partitioned according to the context of its use. In this way, we might have contradicting knowledge in the same knowledge base, but this does not cause any problems.
- *Flexible semantics:* By using context, we can easily choose the appropriate interpretation of possibly ambiguous terms. A good example is the use of the word "glasses": the appropriate meaning would be different in the context of a wine-and-cheese party and in the context of a visit to an ophthalmologist.
- *Flexible Entailment:* Context might effect the entailment relation. For example, in a particular context, entailment might warrant a *closed world assumption* [45] whereas in some other context, this assumption might not be dropped.

Being the largest commonsense knowledge building attempt, CYC [26, 27] has very important pointers on reasoning with an explicit notion of context [24, 25, 32]. Some aspects of the representation of knowledge which are influenced by contextual factors include [25]:

- *Vocabulary:* The vocabulary (i.e., the predicates, functions, and categories) used for representation should be chosen to be appropriate for their intended domain. For example, Mycin and Oncocin [53] overlap significantly in their domains; however Oncocin has some concept of time whereas Mycin does not. This is because these two programs are used for different tasks, thus in different contexts.
- *Granularity and accuracy:* As with the vocabulary, the application area, thus context, determines the granularity and accuracy of the theory.

- *Assumptions:* The assumptions that the task allows often lead to a simplification of the vocabulary. If we try to continue this simplification for large domains, at one point the assumptions become unstable. Thus, either we should use a highly expressive vocabulary or distribute the assumptions to different tasks.

CYC researchers identify two approaches to building large commonsense knowledge bases, and reasoning over them [25].

The straightforward way that a knowledge base builder might choose would be the introduction of an extremely expressive and powerful vocabulary. This approach increases the complexity of the problem, since using such a vocabulary causes difficulties in truth maintenance, and produces a very large search space.

The second way, also the CYC researchers' way, is to make the context dependence of a theory explicit. In this approach, assertions (axioms, statements) are not universally true; they are only true in a context. An assertion in one context might be available for use in a different context by performing a "relative decontextualization"[10].

In reasoning, the ways of using a formal notion of context include the following:

- *A general theory of some topic:* A theory of mechanics, a theory of weather in winter, a theory of what to look for when buying cars, etc. In CYC, such contexts are called the "micro-theories" [26, 24]. Different micro-theories make different assumptions and simplifications about the world. For any topic, there might be different micro-theories of that topic, at varying levels of detail and generality[11].

- *A basis for problem solving:* For some difficult problems, we can form a particular context. We collect all related assumptions, rules, etc. in that context (called the Problem Solving Context (PSC) in CYC [25, 26, 27]), and can process a query (or a group of related queries) in a relatively small search space. Such contexts must be created dynamically and be disposed of afterwards.

- *A context-dependent representation of utterances:* For example, we can use anaphoric and indefinite statements without completely "decontextualizing" them. In this way, for example, the words "the person" might be used in a discourse without identifying him/her exactly.

[10]This process is intuitively the explication of the names of the contexts and constructing a new context which considers assumptions together with their context.

[11]As a technical point, by keeping different theories distinct, the problem of coherence is transformed from that of maintaining global consistency to maintaining local consistency, which in practice is simpler.

Chapter 3

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 [38].) After that introduction, research on the topic was quite silent until the late eighties. McCarthy published his recent ideas on context in [40], a pioneering work by all means. Other notable works on formalizing context are due to Guha [24], Shoham [52], Buvač and Mason [18, 17], and Attardi and Simi [5, 6].

We have reviewed McCarthy's early ideas [38] in Chapter 2. In this chapter, we will review the other logicist formalizations, starting with McCarthy's more recent proposal. The order of the review will be more or less a chronological one: McCarthy; Guha; Shoham; Buvač and Mason; Attardi and Simi.

3.1 McCarthy on Contexts

In his most recent work [40] McCarthy states three reasons for introducing the *formal* notion of context.

First, the use of context allows simple axiomatizations. He exemplifies this by stating that axioms for static blocks world situations can be *lifted*[1] to contexts involving fewer assumptions—to contexts in which the situation changes.

Second, contexts allow us to use a specific vocabulary of and information about a circumstance. An example for this might be the context of a (coded) conversation in which particular terms have particular meanings that they would not have in the language in general. A more concrete use (from Computer Science) can be identified, if we form an analogy with programming language and database concepts. McCarthy's approach might correspond to the use of *local variables* and *local functions* in programming languages, and *views* in database systems. In each case, the meaning of a term depends upon the context in which it is used.

McCarthy's third goal is to propose a mechanism, by which we can build AI systems which are never permanently stuck with the concepts they use at a given time because they can always *transcend* the context they are in. In our view, this brings about two problems:

- *When to transcend a context?*
  Either the system must be smart enough to do so or we must instruct it when to transcend

[1]The term *lifting* is used very frequently in this report. By lifting a predicate (or formula, axiom, etc.) from one context to another, we mean transferring that predicate (or formula, axiom, etc.) to the other context (with appropriate changes, if necessary).

one level up. In the current state-of-the-art, both solutions are quite difficult.

- *Where to transcend?*
McCarthy says [40, p. 1] "Formulas $\texttt{ist}(c,p)$ are always considered as themselves asserted within a context, i.e., we have something like
$\texttt{ist}(c', \texttt{ist}(c,p))$. The regress is infinite but we will show that it is harmless."[2]

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

$$c' : \qquad \texttt{ist}(c,p) \tag{3.1}$$

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

Some properties of context include the following:

1. *Contexts are abstract objects.*
McCarthy says [40, p. 1] "We do not offer a definition [of contexts], but we will offer some examples." Some contexts will be *rich* objects[3] e.g., the situations in Situation Calculus[4]. Some contexts will not be as rich (that is, poor and might be completely described), e.g., some simple micro- theories [24].

2. *Contexts are first class objects.*
We can use contexts in our formulas in the same way we use other objects[5].

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 non-monotonic rule:

$$\forall c_1\ \forall c_2\ \forall p\ \ (c_1 \preceq c_2) \wedge \texttt{ist}(c_1,p) \wedge \neg ab1(c_1,c_2,p) \rightarrow \texttt{ist}(c_2,p)$$

Here, c_2 is a super-context of c_1 and p is a predicate of c_1. (Here, $ab1$ is an abmormality predicate and $\neg ab1(c_1,c_2,p)$ is used to support the non-monotonicity.) In other words, the above rule is a trivial (and basic) lifting rule from a context to its super-context. Interestingly, we can state a similar lifting rule between a context and one of its sub-contexts:

$$\forall c_1\ \forall c_2\ \forall p\ \ (c_1 \preceq c_2) \wedge \texttt{ist}(c_2,p) \wedge \neg ab2(c_1,c_2,p) \rightarrow \texttt{ist}(c_1,p)$$

Note the difference between the abnormality relations ($ab1$ and $ab2$). Intuitively, $ab1$ represents the abnormality in *generalizing* to a super-context, whereas $ab2$ corresponds to the abnormality in *specializing* to a sub-context.

---

[2] McCarthy is not convincing at this point. He, in our view, does not prove that this kind of nesting is harmless; he simply ignores that nesting. In addition to that, do we really need to transcend one level up, that is, to a level which cannot provide us with fruitful information or methods for our problem? How do we stop transcending and decide that our problem cannot be solved within our theory?

[3] A rich object cannot be defined or completely described. A system may be given facts about a rich object but never the complete description [24].

[4] In our conception, Situation Calculus is totally different than Situation Theory and Situation Semantics. More information on Situation Calculus can be found in [36, 28] and on Situation Theory and Situation Semantics in [13, 10, 20].

[5] In our opinion, this is somewhat luxurious for the time being, but having this property will do no harm.

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 the time is specialized to have the value t. We will return to this example in Section 3.1.1.

5. *There are some relations and functions which take contexts as arguments.*
   The function $value(c, t)$ returns the value associated with term t in context c. We will reexamine this function in Section 3.1.1.

6. *Lifting Rules.*
   According to McCarthy [40], the main goal of the use of contexts is to simplify axiomatizations (by allowing us to lift axioms from one context to another). Therefore, one of the properties of context should be its support of the lifting rules. Lifting rules are always asserted in an outer context which should be capable of supporting such rules. We will examine this issue in more detail in Section 3.1.2 and study an example due to McCarthy in Section 5.2.1.

7. *Linguistic vs. Factual Presuppositions.*
   Context might have both linguistic presuppositions and factual presuppositions. An example for a linguistic presupposition might be the demonstratives and indexicals used in an utterance. An example for a factual presupposition might be the objects which occur in a context, i.e., persons, things, etc. [39].

Now, let us give an example on the use of `ist` formulas:

$$c_0: \quad \texttt{ist}(\textit{context-of}(\text{“Sherlock Holmes stories”}), \text{“Holmes is a detective”})$$

asserts that it is true in the context of Sherlock Holmes stories that Holmes is a detective. (The use of English quotations is original with McCarthy; the formal notation is still undecided.) Here, c_0 is considered to be the *outer context*[6].

3.1.1 Relations and Functions Involving Contexts

In [40], McCarthy proposes some functions and relations involving contexts. These are usually employed when we state lifting rules.

- $value(c, t)$ is a function which returns the value of term t in context c:

$$value(\textit{context-of}(\text{“Sherlock Holmes stories”}), \text{“number of wives of Holmes”}) = 0$$

  This states that Holmes has no wife in the context of Sherlock Holmes stories[7].

[6] Notice, on the other hand, that in the context *context-of*(“Sherlock Holmes stories”), Holmes's mother's maiden name does not have a value.

[7] The interpretation of $value(c, t)$ involves a problem that does not arise with $\texttt{ist}(c, p)$, namely, the space in which terms take values may itself be context dependent. McCarthy says [40, p. 2] that many applications would not require this much generality, and this assignment of values to terms might be considered in a fixed domain, i.e., a domain which does not necessitate context dependency.

In our view, we can easily state $value(c, t)$ as an `ist` formula:

$$value(c, t) = \nu \quad \leftrightarrow \quad \texttt{ist}(c, t = \nu)$$

If we accept the above equivalence, the previous problem reduces to context dependence of $=$ in an `ist` formula. This obviously does not warrant any further study, since `ist` formulas were introduced to resolve this problem, i.e.,

- *specialize-time*(t, c) is a context related to c in which the time is specialized to the value t:

$$c_0 : \mathtt{ist}(\textit{specialize-time}(t, c), at(jmc, \textit{Stanford}))$$

  which states that at time t of context c, John McCarthy is at Stanford University.

  Instead of *specialize-time*, it might be convenient to use the predicate *specializes-time*(t, c_1, c_2) and the axiom

$$c_0 : \textit{specializes-time}(t, c_1, c_2) \wedge \mathtt{ist}(c_1, p) \rightarrow \mathtt{ist}(\textit{specialize-time}(t, c1), p)$$

  in which, via *specializes-time*, context c_1 specializes to c_2 at time t. (Thus, *specialize-time*$(t, c_1) = c_2$.) The above axiom relates the function *specialize-time* to the predicate *specializes-time*. The introduction of the new predicate allows us to state lifting rules that have to do with time.

  Instead of specializing on time, we can specialize on *location*, *speaker*, *situation*, *subject matter*, and so on.

- *assuming*(p, c) is another context like context c in which predicate p is assumed (in the natural deduction sense). Using this function, we might dynamically create a context containing the axioms that we desire.

3.1.2 Lifting and Other Advanced Issues

The main motivation of McCarthy in introducing contexts as formal objects was to simplify axioms and increase their generality by using them in various contexts. In his account, *lifting rules* transfer axioms from one context to another and are the only way of relating a context to another. 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 in an axiom, 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:

$$\begin{aligned} c_0 : \forall x \forall y\ & \mathtt{ist}(\textit{above-theory}, on(x, y)) \\ & \rightarrow \mathtt{ist}(\textit{general-blocks-world}, on(x, y)) \end{aligned}$$

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 the *on* predicate 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[8]. The lifting rule is the following:

$$c_0 : \forall x \forall y \forall s\ \mathtt{ist}(\textit{above-theory}, on(x, y)) \rightarrow \mathtt{ist}(\textit{context-of(s)}, on(x, y, s))$$

   where *context-of* is a function returning the context associated with the situation s in which the usual *above-theory* axioms hold.

the context dependence of predicates, in the first place. However, it might be thought that the reduction of *value* to `ist` results in a loss of expressive power of the lifting rules. But, this is not the case either: the expressive power of lifting rules is not lessened. (This fact is proved by Buvač and Mason [18], together with some other mathematical properties of `ist` formulas.)

[8]Once again, here the word "situation" is used in the Situation Calculus sense.

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 change *on* to *üzerinde*, when we move from *above-theory* to *turkish-above-theory*[9]:

$$c_0 : \forall x \forall y \; \texttt{ist}(\textit{above-theory}, \textit{on}(x,y)) \\ \rightarrow \texttt{ist}(\textit{turkish-above-theory}, \textit{üzerinde}(x,y))$$

The most important property of lifting rules is their non-monotonicity. Without non-monotonicity, we cannot use lifting effectively, i.e., we cannot use the lifting rules in the level of generality that we desire, and might run into difficulties similar to the ones encountered in natural regularities [14, 49], cf. Section 2.2. McCarthy proposes circumscription [37] as a tool to implement non-monotonicity.

When we take contexts in the natural deduction sense (as McCarthy suggested [38]), the operations of *entering* and *leaving* a context might be useful and shorten the proofs involving contexts. In this case, $\texttt{ist}(c, p)$ will be analogous to $c \rightarrow p$, and the operation of entering c can be taken as $\textit{assuming}(p, c)$. Then, entering c and inferring p will be equivalent to $\texttt{ist}(c, p)$ in the outer context.

Relative decontextualization is another issue raised by McCarthy's work. He criticizes Quine's notion of eternal sentences[10], because there is no language in which eternal sentences can be expressed. McCarthy 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 [25].

Another advanced problem in which context might be useful is the notion of *mental states* [40]. 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* [35], 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 beliefs.

3.2 Guha on Contexts

Guha encountered the problem of context while he was working on building a large commonsense knowledge base, namely, CYC [26, 27]. (In the previous chapter, we have reviewed the advantages of contexts in knowledge engineering.)

According to Guha, properties of contexts should be similar to those found in McCarthy [38, 40]. Guha models contexts with micro-theories. Micro-theories are theories of (usually) limited domains. For example, we may collect the basic (naive) knowledge of buying and selling into a set of axioms, and may call it the "Commonsense Micro-Theory of Money" [22]. Intuitively, micro-theories are the context's way of seeing the world, and are considered to have the following two basic properties: (i) there is a set of axioms related to each micro-theory, and (ii) there is a vocabulary which tells us the

[9] 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 [12]. Attardi and Simi [5] 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 result in paradoxes, cf. Section 3.5.

[10] Eternal sentences are introduced in [43], and are assumed to be sentences whose meanings do not depend upon context.

syntax and semantics of each predicate and each function specific to the micro-theory. Similar to McCarthy's conception, micro-theories are interrelated via lifting rules stated in an outer context.

`ist` predicates form the basis of Guha's proposal. Guha first identifies the desirable properties of contexts, using a purely technical approach to the problem (since he had his motivation from CYC). He studies ways of using contexts effectively in reasoning, including the following:

- Contexts might be useful in putting together a set of related axioms. In this way, contexts are used as a means for referring to a group of related assertions (closed under entailment) about which something can be said.
- They can be used as a mechanism for combining different theories. If the assertions in one context were not automatically available in other contexts, the system might as well be a set of disconnected knowledge bases. Therefore, by using lifting rules, different micro-theories may be integrated.
- They might be useful in limiting the scope of a theory.
- Using contexts, we might have multiple models of a task. For example, regarding the task of finding out what to do in case of fire, we may offer different models for a workplace and for a house. In a workplace, the first thing to do may be to take away a file of letters, whereas, in a house, the children must be saved first[11].
- Contexts might be used in natural language understanding and representation.

Being the most important component of his system, lifting is studied by Guha in detail. The basic use of lifting in Guha's proposal is the formation of a problem solving context to transfer the axioms of different micro-theories to a PSC (cf. Section 2.3). In this way, different assertions might be made in different contexts and when solving a problem, the system pulls together information from different contexts by way of lifting axioms.

The most important property of lifting rules is their ability to preserve meaning. For example, lifting rules might be used to transfer facts from a (source) context to another (target) context. In the target context, the scope of quantifiers, the interpretation of objects, and even the vocabulary may change. Therefore, when we state a lifting rule, we must take all the possible outcomes into account. In the case of natural language, the problem becomes more complicated since indexicals come into play.

Lifting rules should be definitely non-monotonic as we stated in the previous section. Guha uses default reasoning [26, 24, 25] in the statement of lifting rules. Guha's intuitions about the general lifting rules might be collected into three categories:

- *Default Coreference:* Although there will be differences among contexts, it can be expected that there will be significant similarities and overlaps. As a result, a significant number of terms in different contexts refer to (mean) the same thing. Thus, such terms can be lifted from one context to another without any modification. Similar to terms, we can expect a significant overlap in many formulas, which may be lifted from one context to another without any change. Therefore, it will be a great simplification if we assume that a lifting operation will not require any modification, unless it is explicitly stated that there should be a change.

[11]Clearly, if there were children in the workplace, they would surely be more important. However, we are assuming that there are no chilren in the workplace.

- *Compositional Lifting:* Between contexts, there might be differences in vocabularies (both in the words used, and in the intended denotations of these words). In this case, specifying lifting rules for individual predicates (and functions) should be enough for the system to use these rules in the lifting of formulas involving these predicates. For example, the following lifting rule
$$\forall x\ \texttt{ist}(c_1, tall(x)) \rightarrow \texttt{ist}(c_2, tall(x, Person))$$
should be enough for the system to lift from
$$\texttt{ist}(c_1, tall(A) \wedge tall(B))$$
to
$$\texttt{ist}(c_2, tall(A, Person) \wedge tall(B, Person))$$

- *Coherence:* Since Guha uses lifting in a completely syntactic sense, there is a need for coherence maintenance. For example, in context c_1, *tall* might be a unary predicate, and in contexts c_2 and c_3, it might be a binary predicate such that in c_2, the second argument is a parameter for a population (e.g., $tall(Fred, Soccer\text{-}Player)$), and in c_3, the second argument is a quality (e.g., $tall(Fred, Very)$). In this case, the lifting rules must be stated in a way that the senses of *tall* should not be mixed up in another context, say c_4. Therefore, when lifting different axioms involving a certain predicate from one context to another, we have to ensure that the lifted forms of the axioms use this predicate coherently and in the same sense.

We will return to these intuitions when we present our proposal in Chapter 5.

In the formalization phase, Guha gives the syntax and semantics of the logic of contexts. In the syntax part, he introduces a set of (rich) objects called contexts, a set of constant symbols, etc., and extends the first order logic to use contexts. The most important part is the construction of a context structure. A *context structure* intuitively defines a context's way of describing the world [24, p. 20]. After an outline of syntax, Guha describes the proof theory for his proposal and gives some examples. The issue of lifting is analyzed and it is shown that the proposal is capable of doing the essentials of lifting.

After the formalization, Guha gives examples of lifting and shows ways of using context in building and reasoning with a large knowledge base. The examples are valuable to demonstrate the practical aspects of contexts, however, we will not present them since they are concerned with a rather large "Car Selection" module of CYC and there are many related micro-theories in each example. The reader may refer to Guha's thesis [24, pp. 67–140 and 165–178] for details.

Guha's proposal is rather similar in spirit to that of McCarthy except that it is motivated from a real problem, namely CYC, and works fine in this target domain. Although the proposal accommodates any level of nesting on context, in CYC there are basically two levels: (i) micro-theories, and (ii) the default outer level. The lifting rules and general facts are stated in the outer level, and the problem is solved by the construction of PSC under this level, unless the problem is local to a micro-theory.

3.3 Buvač and Mason on Contexts

Buvač and Mason [18] (and in a more recent work, Buvač, Buvač, and Mason [17]) approach context from a purely mathematical viewpoint. They investigate the simple logical properties of contexts. They also use $\texttt{ist}(c, p)$ to denote context-dependent truth. Using this modality, they extend the classical propositional logic to what they call the *propositional logic of context.* In their proposal, each context is considered to have its own *vocabulary*—a set of propositional atoms which are defined (or meaningful) in that context.

Buvač and Mason discuss the syntax and semantics of a general propositional language of context, and give a Hilbert-style proof system for this language. Their main results are the soundness and completeness proofs of this system. They also provide soundness and completeness results for various extensions of the general system, and prove that their logic is decidable.

Their system has the following two features [18]:

1. A context is modeled by a set of partial truth assignments which describe the possible states of affairs of that context. Then, the `ist` modality is interpreted as validity: $\mathtt{ist}(c, p)$ is true iff the propositional atom p is true in all the truth assignments associated with context c. Since defining `ist` in terms of validity rather than truth leads to a more general system, Buvač and Mason did it in this way. In their system, `ist` is interpreted as truth obtained by placing simple restrictions on the definition of a model, and enriching the set of axioms.

2. The nature of particular contexts is itself context dependent. The example of Buvač and Mason for this is the context of Tweety, which has different interpretations when it is considered in a non-monotonic reasoning literature context, and when it is considered in the context of Tweety & Sylvester (a popular cartoon). This property leads us to consider a context as a sequence of individual contexts rather than a solitary context. In Buvač and Mason's terminology this property is known as *non-flatness* of the system. The acceptance of a sequence of contexts respects the intuition that what holds in a particular context can depend on how this context is reached.

We will not go any further into the details of their system but just note the extensions related to McCarthy's work:

- Lifting is mathematically analyzed and used, rather than just to exemplify the issue. Buvač and Mason advance a way of stating lifting rules so that a fact from one context might be used in another context.

- Although McCarthy does not offer, in our view, a satisfactory mathematical account of why there is no outermost context, Buvač and Mason show that the acceptance of the outermost context simplifies the meta-mathematics of the contexts. They first assume that there is no outermost context and build a proof system on this assumption. Then, they show that introducing the outermost context only simplifies the way they are dealing with non-flatness.

3.4 Shoham on Contexts

Shoham [52] uses the alternative notation p^c to denote that predicate p holds in context c. According to Shoham, every proposition is meaningful in every context [52, p. 400], but the same proposition might have different truth values in different contexts. Thus, his approach is quite different compared to the approaches of McCarthy, Guha, and Buvač and Mason.

Shoham describes a propositional language depending on his *more general than* relation ($\dot{\supset}$). The relation defines a weak partial ordering between contexts; not every pair of contexts are comparable under it. The most important question related to $\dot{\supset}$ is that "Is there a most general (or most specific) context?" Mathematically this corresponds to the question "Is there an upper (or lower) bound on $\dot{\supset}$?" In Shoham's proposal, the question is not answered, but when the system is analyzed the existence of the most general and the most specific contexts is considered.

The language Shoham describes is quite similar to that of the FOL but his relations $\dot{\supset}, \dot{\vee}, \dot{\wedge}$, and $\dot{\neg}$ work over contexts. Here, $x \dot{\wedge} y$ is defined as the greatest lower bound on x and y with respect to $\dot{\supset}$ (if it exists). Similarly, $x \dot{\vee} y$ is defined as a least upper bound of the contexts x and y (if it exists). When defined, $\dot{\neg} x$ is the context which is not comparable[12] to x under $\dot{\supset}$. A context set is *and-closed* if it is closed under conjunction, *or-closed* if it is closed under disjunction, *and-or-closed* if it is both, *not-closed* if it is closed under negation, and simply *closed* if it is all three. From these definitions, we see that if an or-closed context set contains both x and $\dot{\neg} x$ for some x, then the context set contains the most general context, i.e., the *tautological* context. Similarly, under the same condition, an and-closed context set contains the most specific context, i.e., the *contradictory* context.

After the identification of the above properties of his system, Shoham compares the language of contexts with the language of propositions. Drawing on an intuitive analogy between a context c and a predicate *current-context-is(c)*, Shoham identifies the set of contexts with the set of propositions. Using this identification, he defines truth in a context, i.e., p^c, in terms of conventional implication, i.e., *current-context-is(c)* $\rightarrow p$. Then, using a dozen or so benchmark sentences, Shoham characterizes this implication, and points out an interesting interaction between contexts and modal operators[13].

3.5 Attardi and Simi on Contexts

Attardi and Simi [5, 6] offer a "viewpoint" representation which primarily depends on the view of context in a natural deduction sense. According to Attardi and Simi, contexts are sets of reified sentences of the FOL.

The main purpose of Attardi and Simi [5] is to present a formalization of the notion of *viewpoint* as a construct meant for expressing varieties of relativized truth. The formalization is done in a logic which extends FOL through an axiomatization of provability and with the proper reflection rules.

The basic relation in the formalization is

$$\mathsf{in}('A', vp)$$

where A is a sentence provable from viewpoint vp by means of natural deduction techniques. Viewpoints denote *sets of sentences* which represent the axioms of a theory. Viewpoints are defined as a set of reified meta-level sentences.

Two important points from the paper are as follows:

- Attardi and Simi criticize the approaches of McCarthy [40] and Guha [24] for their support of *lifting rules.* Applying such rules in the reasoning, they exhibit the use of logical properties which subsume those required by Montague [41]. Thus, one can always obtain a paradoxical result in a system of the sort suggested by McCarthy (or Guha). In the formalization of Attardi and Simi, this fact is kept in mind and a meta-level is introduced. Reasoning by using this meta-level over viewpoints and the object level is called *contextual entailment.*

- Since viewpoints are defined to be sets of *reified* sentences, operations between viewpoints are

[12] Here, by the term *not-comparable*, we mean that each context contains some axioms which are not contained in the other context.

[13] In [52], Shoham uses the K (knowledge) modality, but a similar discussion holds for other modalities such as belief, choice, etc.

carried out via meta-level rules, e.g.,

$$\frac{\mathsf{in}('B', vp \cup \{'A'\})}{\mathsf{in}('A \rightarrow B', vp)}$$

This corresponds to the following (implication introduction) in classical logic:

$$\frac{vp \cup \{A\} \vdash B}{vp \vdash (A \rightarrow B)}$$

The effective use of viewpoints in making useful proofs requires us to establish a connection between the meta-level and the object-level rules. The *reflection rules* below to do that:

$$\frac{vp_1 \vdash \mathsf{in}('A', vp_2)}{vp_1 \cup vp_2 \vdash A} \qquad \textit{(reflection)}$$

$$\frac{vp \vdash_C A}{\vdash \mathsf{in}('A', vp)} \qquad \textit{(reification)}$$

The notation $\vdash_C$ stands for "classically derivable" or "derivable without using the reflection rules."

Attardi and Simi cite a wide range of examples using the viewpoints. For instance, using viewpoints the notions of *belief*, *knowledge*, *truth*, and *situation*[14] can be formalized as follows:

- *Belief:* The belief of an agent g is captured by means of in sentences, using $vp(g)$ as the viewpoint corresponding to the set of assumptions of the agent. Therefore,

  $$\mathsf{Bel}(g, A) = \mathsf{in}(A, vp(g))$$

  and, by the reflection rule

  $$\mathsf{in}(A, vp(g)) \rightarrow (vp(g) \rightarrow A)$$

  we can use the beliefs of an agent. When we accept this account, the viewpoint $vp(g)$ corresponds to the set of beliefs of g and inferring a belief A from the beliefs of the agent is done by using the reflection rule.

- *Truth:* Truth is captured as provability in a special theory, viz. Real World (RW). Ideally, everything that is true should be derivable in this theory, and truth can be defined as

  $$\mathsf{True}(A) = \mathsf{in}(A, \mathsf{RW})$$

  When we accept this account of truth, our special theory RW should have the following properties: (i) *consistency:* the real world should be non-contradictory; (ii) *completeness:* anything is either true on not; (iii) *veridicality:* the set of assumptions of the real world are true. Using a similar reflection rule as in the belief part, the facts of the world can be used.

- *Knowledge:* Attardi and Simi view knowledge as true belief:

  $$\begin{aligned} \mathsf{K}(g, A) &= \mathsf{Bel}(g, A) \wedge \mathsf{True}(A) \\ &= \mathsf{in}(A, vp(g)) \wedge \mathsf{in}(A, \mathsf{RW}) \end{aligned}$$

  Clearly, all the properties usually ascribed to knowledge can be derived, e.g.,

  $$\mathsf{K}(g, A) \rightarrow A$$

[14]The term *situation* is used in the Situation Theory sense.

- *Situations:* Attardi and Simi take situations as sets of basic facts [12], and use an approach similar to that of belief. Thus, they define a basic relation

$$\mathsf{Holds}(A, s) = \mathsf{in}(A, vp(s))$$

where $vp(s)$ is the set of facts holding in a situation s.

In our view, their formalizations of these notions are not fully satisfactory. First, the notions are treated in a dangerously syntactic way. Regarding all of the above notions solely in a natural deduction sense might cause loss of some of the philosophical properties of these notions [29]. Our second criticism has to do with their proposal for situations: situations are taken to be sets of sentences [20], but the language and the semantics of situations are completely ignored. After a brief introduction to Situation Theory in the next chapter, this last point will be clearer.

Chapter 4

The Situation Theoretic Approach

4.1 An Introduction to Situation Theory

4.1.1 History

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. The basic ideas of the theory were first introduced by Barwise and Perry in the eighties in their influential book *Situations and Attitudes* [13]. Barwise and Etchemendy [12] brought together ideas from Situation Theory and Aczel's work in set theory [1] to study semantical paradoxes. Barwise [10] applies Situation Theory to a variety of problems, including the semantics of natural language conditionals, the nature of constraints, and the characterization of common knowledge. Devlin gives the most up-to-date version of Situation Theory in his excellent book *Logic and Information* [20].

One of the most notable motivations of the theory is to provide a mathematical theory of meaning. There have been different approaches towards building theories of meaning. Some of these theories emphasized the power of language to classify minds, i.e., the *mental significance of language*, while others focused on the connections between language and the described world, i.e., the *external significance of language*. Barwise and Perry [13] claim that for an expression to have meaning, it should convey information[1]. 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.

Devlin also regards Situation Theory as a theory of information. However, in his book, rather than trying to define information, he investigates the nature of information flow and the mechanisms that give rise to it.

4.1.2 Basic Situation Theory

The two major concepts of Situation Theory are infons and situations. Infons are the basic informational units. They should be considered as discrete items of information. Infons 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).

[1] Clearly, this is possible only if the expressions have a link with the kinds of events they describe and also a link with the states of mind.

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 or individuate. The only definition given at this level is that of the *supports* relation:

s *supports* α (denoted $s \models \alpha$) means that α is an infon that is true of s.

It is desirable to have some computational tools to handle situations. Abstract situations are the mathematical constructs which are more amenable to mathematical manipulation. An abstract situation is defined as a (possibly non-well-founded [1]) set of infons. Given a real situation s, the set $\{\alpha \mid s \models \alpha\}$ is the corresponding abstract situation.

One of the important ideas behind Situation Theory is a *schema of individuation*, a way of carving the world into 'uniformities.' The notions of individuals, relations, spatial and temporal locations, and further entities depend upon this schema of individuation. Thus, the constituents of Situation Theory such as infons, constraints, and situations are determined by the agent's schema of individuation. We will try to follow a rational agent's (probably an agent having a reasonable math background) individuation schema in our examples.

Being constructs that link the schema of individuation to the technical framework of the theory, types are important features of Situation Theory. Types are higher-order uniformities which cut across uniformities such as individuals, relations, situations, and spatio-temporal locations. Just as individuals, temporal locations, spatial locations, relations, and situations, types are also uniformities that are discriminated by agents. Relations may have their argument places filled either with individuals, situations, locations, and other relations, or with types of individuals, situations, locations, and relations.

In Situation Theory, for each type T, an infinite collection of *parameters* $T_1, T_2, \ldots$ is introduced. For example IND_3 is an IND-parameter. These bring about some computational power, but we may demand more than that. Sometimes, rather than parameters ranging over all individuals, we need parameters that range over a more restricted class, viz. *restricted parameters*[2]. For example,

$$\dot{r1} = \dot{a} \upharpoonright \ll kicking, \dot{a}, \dot{b}, 1 \gg$$

$$\dot{a} = IND3 \upharpoonright \ll man, IND3, 1 \gg$$

$$\dot{b} = IND2 \upharpoonright \ll football, IND2, 1 \gg$$

In this case, $\dot{r1}$ ranges over all men kicking footballs.

Related to the parametric infons, there is a construct by which we can assign 'values' to parameters. We call this an *anchor*. 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. Therefore, if f is an anchor for A and T_i is a parameter in A, then

$$\ll \textit{of-type}, f(T_i), T, 1 \gg$$

For example, if f anchors $\dot{a}$ to the type IND_3 individual "Surav," we write

$$f(\dot{a}) = \text{Surav}$$

to denote this anchoring.

[2]we represent restricted parameters with a parameter nama which has a dot over it.

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 with

$$\ll involves, S_0, S_1, 1 \gg$$

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, with the following constraint c, we might represent the regularity "Smoke means fire"[3]:

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

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

$$c = \ll involves, S_0, S_1, 1 \gg$$

Let us return to the relation between types and propositions. Representation of propositions is an important issue for us. In this discussion, the role of the grounding situation becomes more apparent. Basically, in Situation Theory the 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 $f(\dot{x}) = x$. Thus all propositions are essentially of the form

$$< \text{situation} > \models < \text{infon} >$$

[3]Barwise [8] investigates conditional knowledge and its dependence upon circumstantial information. Since conditionals are tightly related to constraints, we will stop the discussion here and continue when we review Barwise's contributions in the next section.

4.2 Barwise on Contexts

Barwise's ideas on circumstance, thus on context, arise from his work on conditionals and circumstantial information [8]. Barwise states two innovative examples, one on a missing pollen and the other on the (wrong) proof of $1 = -1$.

Example 1: 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 general conditional statement

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, Barwise and his wife eventually figured out that Claire was allergic to something or other. They called it Pollen X since they did not know the exact name. So Pollen X could also cause Claire to rub her eyes. □

Example 2: Proof of $1 = -1$
We will give a very simple 'proof' that $1 = -1$, which uses (or rather, misuses) true conditional statements and the usual laws of equality. In the proof, i represents $\sqrt{-1}$, so that $i \times i = -1$. We will also use the fact that $\sqrt{1} = 1$. Now,

$$If \;\; x = a \cdot b, \;\; then \;\; \sqrt{x} = \sqrt{a} \cdot \sqrt{b} \tag{4.1}$$

$$If \;\; x = 1 \;\; and \;\; a = -1 \;\; and \;\; b = -1, \;\; then \;\; x = a \cdot b \tag{4.2}$$

From these true statements, and the laws of equality, we can conclude, using the Hypothetical Syllogism, that

$$If \;\; x = 1 \;\; and \;\; a = -1 \;\; and \;\; b = -1, \;\; then \;\; x = -1. \square$$

Barwise approaches the problems stated in the above examples 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 we can represent the information that "If Claire rubs her eyes, then she is sleepy" with the following parametric constraint C (with unspecified time and location parameters):

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

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

$$C = [S \Rightarrow S']$$

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 deal with the problem:

- 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]

where β corresponds to the additional background conditions.

Barwise chooses the second way to deal with the problem[4] and modifies the *involves* relation, and makes the background assumptions implicit by introducing a third parameter. The *involves* relation now becomes:

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

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

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

Barwise enumerates the following five assumptions that govern the new *involves* relation:

1. If B is fixed, then the resulting two-place relation is transitive, i.e., $S_1 \Rightarrow S_2|B$ and $S_2 \Rightarrow S_3|B$ then $S_1 \Rightarrow S_3|B$. (This is why the Hypothetical Syllogism is valid as long as the background conditions do not vary.)

2. If a conditional constraint holds relative to some B and B is tightened, then the constraint holds relative to the more restrictive type of situation B', i.e., if $S \Rightarrow S'|B$ and $B' \subset B$, where B' is compatible with S, then $S \Rightarrow S'|B'$.

3. if $S \Rightarrow S'|B$ then S is compatible with B, i.e., $S \cap B$ is not incoherent[5].

4. if $S \Rightarrow S'|B$ and f is a coherent parameter anchor for some of the parameters of B, then $S(f) \Rightarrow S'(f)|B(f)$.

5. if $S \Rightarrow S'|B$ where B has no parameters, and if B is realized by some real situation, then $S \Rightarrow S'$ is actual.

These properties form a basis for our notion of context. Our aim in this report is to design our contexts so that the above properties apply to lifting operations between contexts.

Returning to Barwise, the two examples are now revisited with additional light shed by the new *involves* relation.

Example 1: Missing Pollen.
We can reformulate the problem using a background condition (B):

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

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

$$B = [\dot{s}|\dot{s} \models \ll exists, \text{Pollen X}, \dot{l}, \dot{t}, 0 \gg]$$

$$C = [S \Rightarrow S'|B]$$

[4]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 constraint (i.e., if ϕ then ψ) is directly modified to use background conditions, whereas, in the second case, the constraint is not touched, but is evaluated only when background conditions hold. Introduction of background conditions for constraints corresponds to a non-monotonic reasoning mechanism.

[5]Here, the term *incoherent* is used in the sense that $S \cap B$ does not contain any contradictory infons, i.e., $\ll R, ..., 0 \gg$ and $\ll R, ..., 1 \gg$ do not occur together in $S \cap B$ for any R.

In this example, B will be the conditions to be supplied by the grounding situation, and thus by the context. We will review this example once again in Section 5.2.2, and give a context which has this information.

Example 2: Proof of $1 = -1$**.**
In the example, the conditional which depends on the environmental factors, i.e., the constraint effected from the situation shift, is Equation 4.1. This constraint can be modeled by C:

$$S = [\dot{s}|\dot{s} \models \ll product, \dot{a}, \dot{b}, \dot{x}, 1 \gg]$$

$$S' = [\dot{s}|\dot{s} \models \ll product, \sqrt{\dot{a}}, \sqrt{\dot{b}}, \sqrt{\dot{x}}, 1 \gg]$$

$$B = [\dot{s}|\dot{s} \models \ll \text{All numbers are positive real numbers} \gg]$$

$$C = \ll involves, S, S', B, 1 \gg$$

where B might be explicitly represented using C_B:

$$S_B = [\dot{s}|\dot{s} \models \ll number, \dot{a}, 1 \gg]$$

$$S'_B = [\dot{s}|\dot{s} \models \ll real, \dot{a}, 1 \gg \wedge\ \dot{s} \models \ll positive, \dot{a}, 1 \gg]$$

$$C_B = \ll involves, S_B, S'_B, B_B, 1 \gg$$

and

$$B \models C_B$$

Here, B_B is yet another background condition[6].

[6] This is not important in our example, and we will not describe it. However, if one requires to state a background condition for C_B, he can freely do that.

Chapter 5

Our Proposal

In this chapter, we will detail our approach and present a model for context in the framework of Situation Theory. In Section 5.1, we will describe our formalization, and show that it is as powerful as the other approaches in the literature vis-à-vis the desired properties of context. In Section 5.2, we will present the examples that we promised in the previous chapters.

5.1 Motivation

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

1. Offering a representation schema which allows the contexts of Logic, Reasoning, and Natural Language in a uniform way.
2. Defining a representation schema which supports the essential properties of context. These properties were explained in Chapters 3 and 4.
3. Clarifying the notion of context and reducing the problem to a computer science (or mathematics) problem. As future work, the implementation of the proposal might be considered.

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:

- *parameter free* infons to state the facts and the usual bindings.
- *parametric* infons (which correspond to parametric conditionals representing infons to represent 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 Surav'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

$$c \models \ll \textit{school}, \text{Bilkent}, 1 \gg \tag{5.1}$$

$$c \models \ll \textit{department}, \text{Computer Engineering}, 1 \gg \tag{5.2}$$

$$c \models \ll \textit{ms-student}, \text{Surav}, 1 \gg \tag{5.3}$$

$$c \models \ll \textit{ms-advisor}, \text{Akman}, 1 \gg \tag{5.4}$$

$$c \models \ll \textit{ms-jury-member}, \text{ Çiçekli}, 1 \gg \tag{5.5}$$

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

$$S_1 \models \ll \textit{ms-advisor}, \dot{a}, 1 \gg \tag{5.6}$$

$$S_2 \models \ll \textit{ms-jury-member}, \dot{a}, 1 \gg \tag{5.7}$$

$$c_1 = \quad [S_1 \Rightarrow S_2 | B] \tag{5.8}$$

where B is a background situation such as

$$B \models \ll \textit{school}, \text{Bilkent}, 1 \gg \tag{5.9}$$

The constraint c_1 can be represented with the following infon:

$$c_1 \models \ll \textit{involves}, S_1, S_2, B \gg \tag{5.10}$$

The second part, i.e., the set of infons in Equations 5.6–5.10, 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

$$f(\dot{a}) = \text{Akman},$$

we can conclude that Akman is also a jury member.

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

During the review of McCarthy's work (cf. Section 3.1), the very first property we stated was 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 [38, 40] and Guha [24]. In Situation Theory, situations are, by definition, rich objects [20, 12] (cf. Section 3.1 for the definition of rich object.). The richness of situations covers the partiality of contexts as McCarthy desires.

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 contain no variables (either bound or unbound) and rules were represented with quantified logical implications:

$$c: \quad \textit{present}(Air) \tag{5.11}$$

$$c: \quad \forall x \; bird(x) \rightarrow flies(x) \tag{5.12}$$

Equation 5.11 states that air is present (a presupposition), and Equation 5.12 states that if something is a bird, it flies (a default rule). The same capability is also available in our notion of context. In Situation Theory, 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 examples might then be restated as the following situation-theoretic constructs:

$$S_c \models \ll present, \mathrm{Air}, 1 \gg \tag{5.13}$$

$$\begin{array}{l} S_1 = [\dot{s} | \dot{s} \models \ll bird, \dot{a}, 1 \gg \\ S_2 = [\dot{s} | \dot{s} \models \ll flies, \dot{a}, 1 \gg \\ S_c \models \ll involves, S_1, S_2, 1 \gg \end{array} \tag{5.14}$$

Here, the fact that air is present is represented with the infon in Equation 5.13, and the rule that birds fly is represented with the constraint in Equation 5.14. Therefore, we can use the situation type S_c as as the context of the logicist approach (namely, context c).

Another way of using the context is to fill in the missing parameters of statements in natural language[1]. We will not consider transferring natural language statements to logic, but only present the way we can use it in our model. Since we will give an example in Section 5.2.5, we will only study why our model is capable of realizing the appropriate parameter filling task, for the moment.

A related issue is the background information. In his paper "Conditionals and Conditional Information" [8], Barwise points out the importance of the background information in the *involves* ($\Rightarrow$) relation. In our model, the context representation is designed to supply adequate background information.

Consider the context of M.S. Thesis presentation stated in the beginning of this chapter. In this context, the advisor of the thesis is also a jury member of the thesis. This rule is stated formally via c_1 in Equation 5.8. In the constraint, the background condition is that the school is Bilkent University. This is stated by using B of Equation 5.9. In our definition of context, supplying this kind of background information is simple (and in fact necessary). We will give an example in Section 5.2.5.

Another natural language construct related to context is the indexical. We have promised in Chapter 2 (cf. the example "I am a philosopher") that our notion of context will be capable of dealing with indexicals. What we do is an adaptation of the previous approaches to our notion of context. This is not difficult, since our contexts are, in some sense, situations. When we review the example "I am a philosopher" in Section 5.2.5, we will demonstrate how we use our model for the task of matching indexicals.

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, we can use some axioms from one context in another context. The issue of lifting is raised in the logicist approach. Lifting rules (whether non-monotonic or not) are always stated in

[1]The task of filling the missing parameters is in fact a matching between the objects available in the context and the missing parameters. To simplify the problem, the objects of the context might be partitioned into particular types. (With this partitioning, we might reduce the search space.) In the case of ordinary (unsorted) FOL, since we have no type casting over the objects, the search space is large and since matching is an NP-complete problem, doing it is a time consuming process. However, in Situation Theory, since the objects are type cast, this process becomes simpler and can effectively be used in practice. (The example "Carl Lewis is running." of Section 5.2.5 corresponds to this.) Note that, finding the appropriate anchoring function is still a difficult task.

the outer one of the two contexts between which the lifting will be done. In our case, lifting has similar properties. Lifting rules are stated in an outer context and are non-monotonic.

Basically, we will state lifting rules as constraints. Non-monotonicity 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:

$$S_1 = [\dot{s} | \; C_1 \models \ll foo, \dot{a}, 1 \gg] \tag{5.15}$$

$$S_2 = [\dot{s} | \; C_2 \models \ll bar, \dot{a}, 1 \gg] \tag{5.16}$$

$$C \models \ll involves, S_1, S_2, B, 1 \gg \tag{5.17}$$

B is the background condition, which enables us to have non-monotonicity 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 due to the philosophy behind Situation Theory. In Situation Theory we use an individuation mechanism to name objects, individuals, events, situations, and so on. One application area of the individuation mechanism is relations. For example, if we individuate the *on* relation with two parameters as

$$\ll on, \dot{a}, \dot{b}, 1 \gg \tag{5.18}$$

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.

Regarding Guha's discussions on lifting, the intuitions *default coreference, compositionality,* and *coherence* should also be guaranteed (cf. Section 3.2). In the framework of Situation Theory, coherence is not a problem, since the meanings of relations are determined by the schema of individuation. As for the compositionality, the basic unit of information in Situation Theory is the infon, and constraints are the way to the building of formulas. In the lifting of formulas, it is not enough for us to lift only infons; we also require the lifting of constraints. Finally, default coreference does not apply to Situation Theory.

5.2 Examples

After the statement of the properties of our formal context, we now show how it works for the various examples from the literature. In previous chapters, we have reviewed the relevant works, but have left the analysis of the examples to this chapter. We will begin with an example due to McCarthy. This is rather longish but is the most significant outcome of our proposal, since it creates a parallelism between predicate calculus and Situation Theory. In the remaining examples, we will show that our notion of context is adequate to deal with various kinds of context-dependent problems (from natural language, logic, categorization, etc.).

5.2.1 McCarthy's Lifting Example

In [40], McCarthy states the following example in his "Lifting Rules" section. Here, we will first present the original example, and then re-do it in our version of formal context.

McCarthy uses the blocks world to state his example.

Original Lifting Example.

McCarthy considers two contexts, namely, *Above-Theory (AT)* and c. *Above-Theory* is the context which contains some simple blocks-world assumptions, similar to Equations 5.19–5.20. In AT, the notion of situation is undefined. However, the context 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 (in some sense a sub-context) of c, where the situation is fixed to s. The lifting rules working between c and one of its specializations are presented as Equations 5.21 and 5.22. Equation 5.23 is the major lifting axiom, which links AT and the sub-contexts of c. The example in McCarthy [40] is the proof that from $\texttt{ist}(c, on(A, B, S_0))$ we can prove that $\texttt{ist}(c, above(A, B, S_0))$. In the proof, c_0 is the outer context. The axioms are the following[2]:

$$c_0 :\ AT :\ \forall x\, \forall y\ \ on(x, y) \rightarrow above(x, y) \tag{5.19}$$

$$c_0 :\ AT :\ \forall x\, \forall y\, \forall z\ \ above(x, y) \wedge above(y, z) \rightarrow above(x, z) \tag{5.20}$$

$$c_0 :\ c :\ \forall x\, \forall y\, \forall s\ \ on(x, y, s) \leftrightarrow \texttt{ist}(\textit{context-of}(s), on(x, y)) \tag{5.21}$$

$$c_0 :\ c :\ \forall x\, \forall y\, \forall s\ \ above(x, y, s) \leftrightarrow \texttt{ist}(\textit{context-of}(s), above(x, y)) \tag{5.22}$$

$$c_0 :\ c :\ \forall p\, \forall s\ \ \texttt{ist}(AT, p) \rightarrow \texttt{ist}(\textit{context-of}(s), p) \tag{5.23}$$

The proof proceeds as follows:

$$c_0 :\ c :\ on(A, B, S_0) \tag{5.24}$$

$$c_0 :\ c :\ \texttt{ist}(\textit{context-of}(S_0), on(A, B)) \tag{5.25}$$

$$c_0 :\ c :\ \textit{context-of}(S_0) :\ on(A, B) \tag{5.26}$$

$$c_0 :\ c :\ \texttt{ist}(\textit{context-of}(S_0), \forall x\, \forall y\ on(x, y) \rightarrow above(x, y)) \tag{5.27}$$

$$c_0 :\ c :\ \textit{context-of}(S_0) :\ above(A, B) \tag{5.28}$$

$$c_0 :\ c :\ above(A, B, S_0) \tag{5.29}$$

Equation 5.24 is the assumption given in the beginning. Equation 5.25 is obtained from Equations 5.21 and 5.24 by binding A to x, B to y, and S_0 to s. Equation 5.26 obtained from Equation 5.25 by entering the context *context-of*(S_0). Equation 5.27 is the result of lifting Equation 5.19 by

[2] As a reminder, the formula $\texttt{ist}(c, p)$ can also be represented as $c : p$. This states that predicate p holds in context c.

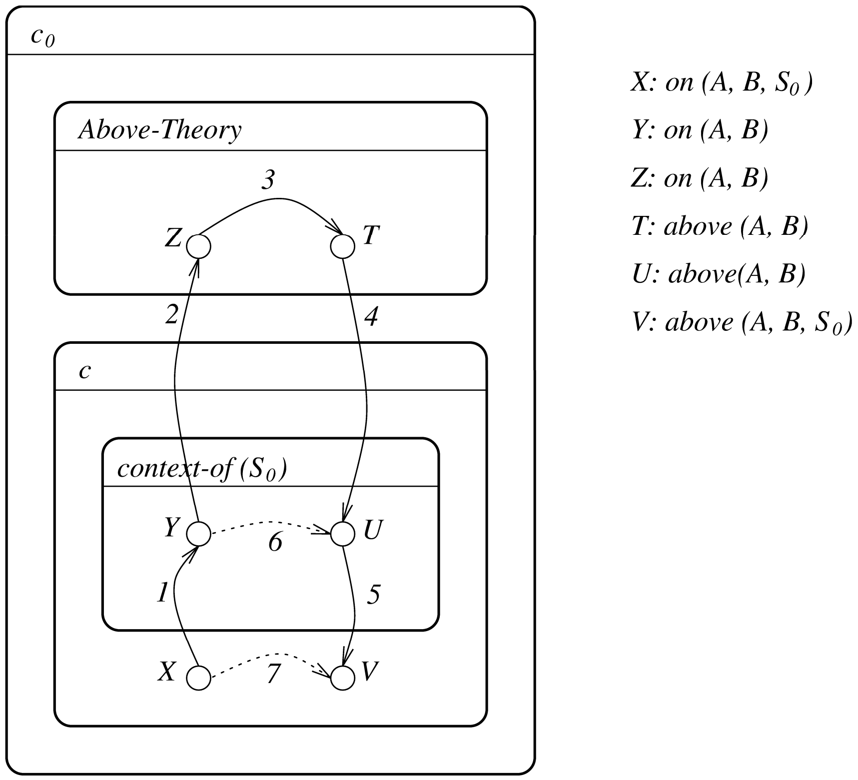

Figure 5.1: Diagram of McCarthy's proof

the lifting axiom (Equation 5.23). From Equations 5.26 and 5.27, we obtain Equation 5.28. The desired equation 5.29 is obtained from Equations 5.28 and 5.22. This proof is summarized in Figure 5.1. In the figure, 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.' If we look at the proof from a conceptual level, McCarthy is drawing a virtual link from the atomic formula X to the atomic formula V. Since, c has no rule to draw 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 5.21. 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 can 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's work [6], this link is explicitly stated and a proof is carried out with the path 1-2-3-4-5.

We can rework this example within our framework. 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 5.19) 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[3].

- *context-of*(S_0) function of McCarthy will be represented with the infons of type

$$\ll context\text{-}of, A, B, 1 \gg$$

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.)

- The background conditions $B_{AT}, B_{c\text{-}AT}, B_{AT\text{-}c}$[4] will not be explicitly stated in the proof, since the original proof does not involve any non-monotonic inference.

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

$$\begin{array}{l} S_{11} = [\dot{s} | \dot{s} \models \ll on_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ S_{12} = [\dot{s} | \dot{s} \models \ll above_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ c_{AT} \models \ll involves, S_{11}, S_{12}, B_{AT} \gg \end{array} \tag{5.30}$$

$$\begin{array}{l} S_{21} = [\dot{s} | \dot{s} \models \ll above_{AT}, \dot{x}, \dot{y}, 1 \gg \wedge \ \dot{s} \models \ll above_{AT}, \dot{y}, \dot{z}, 1 \gg] \\ S_{22} = [\dot{s} | \dot{s} \models \ll above_{AT}, \dot{x}, \dot{z}, 1 \gg] \\ c_{AT} \models \ll involves, S_{21}, S_{22}, B_{AT} \gg \end{array} \tag{5.31}$$

$$\begin{array}{l} S_{31} = [\dot{s} | \dot{s} \models \ll on_c, \dot{x}, \dot{y}, \dot{s_{31}}, 1 \gg] \\ c_{AT} = [\dot{s} | \dot{s} \models \ll on_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ c_c \models \ll context\text{-}of, \dot{s}, \dot{s_{31}}, 1 \gg \\ c_c \models \ll involves, S_{31}, c_{AT}, B_{c\text{-}AT} \gg \end{array} \tag{5.32}$$

$$\begin{array}{l} S_{31} = [\dot{s} | \dot{s} \models \ll above_c, \dot{x}, \dot{y}, \dot{s_{41}}, 1 \gg] \\ c_{AT} = [\dot{s} | \dot{s} \models \ll above_{AT}, \dot{x}, \dot{y}, 1 \gg] \\ c_c \models \ll context\text{-}of, \dot{s}, \dot{s_{41}}, 1 \gg \\ c_c \models \ll involves, S_{41}, c_{AT}, B_{c\text{-}AT} \gg \end{array} \tag{5.33}$$

$$\begin{array}{l} c_{AT} \models \sigma_p \\ c_c \models \ll context\text{-}of, \dot{s}, \dot{s_{51}}, 1 \gg \\ \dot{s_{51}} \models \sigma_p \\ c_c \models \ll involves, c_{AT}, \dot{s_{51}}, B_{c\text{-}AT} \gg \end{array} \tag{5.34}$$

In addition to McCarthy's axioms, we will need a further rule to lift facts from c_{AT} to c_c. This is the following constraint:

[3]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 (cf. discussion on schema of individuation in Section 4.1.2).

[4]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.

$$
\begin{array}{l}
S_{61} = [\dot{s}|\dot{s} \models \ll \mathit{above}_c, \dot{x}, \dot{y}, \dot{s_{61}}, 1 \gg] \\
c_{AT} = [\dot{s}|\dot{s} \models \ll \mathit{above}_{AT}, \dot{x}, \dot{y}, 1 \gg] \\
c_c \models \ll \mathit{context\text{-}of}, \dot{s}, \dot{s_{61}}, 1 \gg \\
c_c \models \ll \mathit{involves}, c_{AT}, S_{61}, B_{AT\text{-}c} \gg
\end{array}
\tag{5.35}
$$

In McCarthy, we have the following

$$c_0 : \mathtt{ist}(c, \mathit{on}(A, B, S_0)). \tag{5.36}$$

This can be represented with the following infon

$$\sigma_0 = \ll \mathit{on}_c, A, B, S_0, 1 \gg$$

and then Equation 5.36 corresponds to the following:

$$c_c \models \sigma_0 \tag{5.37}$$

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

$$c_c \models \ll \mathit{above}_c, A, B, S_0, 1 \gg \tag{5.38}$$

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 5.1.

Using the constraint of Equation 5.32 with the following anchoring

$$f_1(\dot{s}) = S_0$$

$$f_1(\dot{x}) = A$$

$$f_1(\dot{y}) = B$$

we transfer $\ll \mathit{on}_c, A, B, S_0, 1 \gg$ from c_c to $\ll \mathit{on}_{AT}, A, B, 1 \gg$ in c_{AT}. This corresponds to tracing links 1 and 2 in Figure 5.1. 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 following anchoring

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

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

and Equation 5.30, we get $\ll \mathit{above}_{AT}, A, B, 1 \gg$. This corresponds to link 3 in Figure 5.1. After this implication of *above* from *on*, we should transfer the fact to c_c. This is done using Equation 5.35 with f_1. The result of this operation is $\ll \mathit{above}_c, A, B, S_0, 1 \gg$. This completes the proof path 1-2-3-4-5 in Figure 5.1. Once more, by using one constraint, we have traced two links, namely, 4 and 5, in Figure 5.1. 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.

5.2.2 Examples from Barwise

Missing Pollen.

The following are the constituents of the constraint C, which was the solution to the missing pollen problem:

$$S = [\dot{s}|\dot{s} \models \ll rubs, \text{Claire, Claire's eyes}, \dot{l}, \dot{t}, 1 \gg] \tag{5.39}$$

$$S' = [\dot{s}|\dot{s} \models \ll sleepy, \text{Claire}, \dot{l}, \dot{t}, 1 \gg] \tag{5.40}$$

$$B = [\dot{s}|\dot{s} \models \ll exists, \text{Pollen X}, \dot{l}, \dot{t}, 0 \gg] \tag{5.41}$$

$$C = \ll involves, S, S', B, 1 \gg \tag{5.42}$$

At the beginning it was winter and there were no pollens. The context of the talk, 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[5] the background condition B (Equation 5.41) of constraint C (Equation 5.42), and thus can conclude that "Claire is sleepy."

Later, in summer, the new context, call it 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" becomes cannot be reached/proved.

$1 = -1$ Example.

This is similar to the above example. In the (wrong) proof of $1 = -1$, there exists a conflict between the context of the proof and the background condition of the constraint, which leads to this false equality.

5.2.3 Perspectives Example

In Chapter 2, we pointed out that interpretations of some words directly depend upon context. In Figure 2.1, we have used the following sentence to exemplify this:

Engineering Building is to the left of the Library.

Similar issues are raised by Seligman and Barwise [48, 14, 49] who present more mathematical approaches based on the flow of information [21]. In addition to the above works, our account of context may also be used to present a satisfactory way of modeling these issues[6].

[5] Since we are using background conditions as a tool for non-monotonicity, we are looking for the opposites of the background conditions to appear in the context. If we do not find any opposite in the context, we conclude that background conditions are not being violated.

[6] Our approach does not analyze the problem, but offers an *ad hoc* solution.

Let A be an agent looking towards the Engineering Building from the Tourism School, and B be another agent looking from the Publishing Company. Let c_A and c_B denote the contexts of A and B, respectively. In c_A, we have the following information:

$$\ll \textit{left-of}, \text{ Engineering Building, Library,1} \gg \tag{5.43}$$

$$\ll \textit{looking-from}, \text{ Tourism School,1} \gg \tag{5.44}$$

In c_B, we have

$$\ll \textit{right-of}, \text{ Engineering Building, Library,1} \gg \tag{5.45}$$

$$\ll \textit{looking-from}, \text{ Publishing Company,1} \gg \tag{5.46}$$

Now, the relations *left-of* and *right-of* become non-problematic, since the context will be significant as a grounding situation in interpretations of sentences. Thus, the utterance "Engineering Building is to the left of the Library" is supported in c_A, but is not supported in c_B.

5.2.4 The Springfield Example

In the Springfield barbers example [2, 55], our solution was to introduce of a context-dependent membership. However, with the availability of a formal notion of context, we do not need any additional construct, or modification in the membership relation.

In the example, the number of barbers at Springfield was different when looked at it from a commonsense point of view and from a legal point of view. Both of the views corresponds to different contexts, and if we accept this, we can still be consistent with these different numbers. In the statement of the example, there was a barber who does not work for money, but serves the community by cutting their hair and shaving them. Let this person be John, the librarian of the town.

In the context of the commonsense view, call it c_{CS}, we have the following

$$c_{CS} \models \ll barber, John, 1 \gg$$

$$c_{CS} \models \ll librarian, John, 1 \gg$$

In the government agencies' view, the context of the talk, call it c_{GA}, will only have the following

$$c_{GA} \models \ll librarian, John, 1 \gg$$

In this case, when one does a counting from a commonsense point of view, he uses c_{CS} as the grounding situation. In the case of the government agencies view, c_{GA} will be used as the grounding situation in the counting, and the number of barbers will be one less than the number of barbers obtained by the commonsense view.

5.2.5 Natural Language Examples

In this section, we will review some of the natural language examples we have introduced in the previous sections of this report. The only new one is the final example which is a typical non-monotonic reasoning example.

I am a philosopher.

The first point of our discussion was the context dependency of the word "philosopher" (cf. Chapter 2). Since, the conversation has no fruitful pointers to extract the meaning of the word philosopher, we will not discuss the meaning of it, but we will try to prove another fact: the content of all three sentences are the same, i.e., A is a philosopher. In dealing with this issue, we will also study the use of indexicals in our model.

We might have three contexts associated with each individual in the conversation. We will call these contexts c_A, c_B, and c_C, respectively. We will represent the indexicals with special parameters $\dot{I}, \dot{You}$, and $\dot{She}$ which correspond to I, you, and she, respectively.

In c_A, we have the following infons supported:

$$\ll corresponds, \dot{I}, A, 1 \gg$$

$$\ll philosopher, \dot{I}, 1 \gg$$

where *corresponds* is a function which associates an indexical to a person, and utterances about being a philosopher are represented with infons of type $\ll philosopher, \dot{x}, 1 \gg$.

c_B supports the following:

$$\ll corresponds, \dot{She}, A, 1 \gg$$

$$\ll philosopher, \dot{She}, 1 \gg$$

c_C supports

$$\ll corresponds, \dot{You}, A, 1 \gg$$

$$\ll philosopher, \dot{You}, 1 \gg$$

The problem is to find the anchoring function. In this case, we have to believe that the above conversation is correctly transcribed, so that who is talking to whom is, by assumption, correctly stated. Now, it is a trivial matter to observe that $\dot{I}, \dot{You}$, and $\dot{She}$ all collapse to A, i.e., the anchoring

$$f(\dot{I}) = A$$

$$f(\dot{You}) = B$$

$$f(\dot{She}) = C$$

Consequently, the utterance of A might be de-contextualized as

$$\ll philosopher, A, 1 \gg \tag{5.47}$$

Lewis is Running.

In Chapter 2, we have stated that context might be used to fill the missing parameters of some actions in natural language utterances. For example, in

Carl Lewis is running

we can fill the place and time of the action if we know that we are watching Lewis on TV at the 1992 Barcelona Olympic Games. Let c be the context of the above talk which supports the following infons:

$$\ll \textit{championship}, \text{Olympic Games}, 1 \gg \tag{5.48}$$

$$\ll \textit{championship-place}, \text{Barcelona}, 1 \gg \tag{5.49}$$

$$\ll \textit{championship-time}, 1992, 1 \gg \tag{5.50}$$

Using the commonsense rule "a running action takes place in the location of the championship," we can fill the missing parameters of *running* with the place and time of the championship available in c. Thus, we get

$$\ll \textit{running}, \text{Lewis, Barcelona, 1992,1} \gg \tag{5.51}$$

Non-monotonicity.

In Equation 5.12, if x is somehow bound to a non-flying bird like a penguin, this implication is invalidated using some non-monotonic techniques (e.g., circumscription [37], default reasoning [46], etc.).

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

$$S_1 = [\dot{s} | \dot{s} \models \ll \textit{bird}, \dot{x}, 1 \gg] \tag{5.52}$$

$$S_2 = [\dot{s} | \dot{s} \models \ll \textit{flies}, \dot{x}, 1 \gg] \tag{5.53}$$

$$\begin{aligned} B = [\dot{s} | \quad & \dot{s} \models \ll \textit{penguin}, \dot{x}, 0 \gg \\ \wedge \quad & \dot{s} \models \ll \textit{present}, \text{Air}, 1 \gg] \end{aligned} \tag{5.54}$$

$$C = \ll \textit{involves}, S_1, S_2, B, 1 \gg \tag{5.55}$$

The constraint 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. Therefore, within an appropriately defined context, these kind of constraints can be effectively used.

Chapter 6

Conclusion

In this report, we have proposed a formalization of context using Situation Theory. Although we have not stated a specific area of application, the new formalism might be useful for Natural Language Semantics, Reasoning in AI, and Categorization tasks (possibly with some slight revisions). The purpose of our formalization was to build a mathematical basis for the notion of context, which can effectively be used in the above areas.

In the literature, there are number of attempts towards a formalization of context in a logicist framework. Our approach differs from most of these approaches on being stated in the framework of Situation Theory. Our approach is primarily an extension of Barwise's conception of context. In [8], Barwise uses grounding situations similar to our contexts. However, in his work, the 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 relations (cf. Section 5.1 for full definition). From Barwise's work, we could not get any intuition of particular in-context rules.

The comparison of the previous approaches and our approach is summarized in Table 6.1 where the first line categorizes the language of statement. Except for Barwise [8], all of the previous approaches were stated in a more or less logicist framework. Among these, only Shoham proposes context as a modal operator; the other logicists consider context in a natural deduction sense, and allow operations of entering/leaving contexts.

Among the previous approaches, McCarthy's and Guha's are not paradox free, whereas Buvač and Mason's, and Attardi and Simi's approaches are paradox free. We do not know whether Barwise's, Shoham's, and our approaches are paradox free or not. However, in a thought-provoking work [12], Situation Theory is shown to be powerful enough to deal with circularity.

Compared to other approaches, our approach has the following notable properties (some advantageous, some not):

- *Allowing dynamic contexts:* In our approach, 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 approaches, since one can always add (delete) axioms into (from) a theory. However, when we fortify our context with dynamic constraints whose background conditions are dynamic, we get non-monotonicity in the framework of Situation Theory (cf. "A non-monotonic reasoning example" from Section 5.2).

- *Simpler natural language interfaces:* Situation Theory supports a more natural outlook regarding natural language concepts. Thus, our approach might lead to a simpler interface in

	Mc87	Mc93	Gu91	Ba86	Sh91	BM93	AS93	Ours
Logic vs. Situation Theory	Logic	Logic	Logic	S.T.	Logic	Logic	Logic	S.T.
Modal Treatment	No	No	No	No	Yes	No	No	No
Natural Deduction	Yes	Yes	Yes	No	No	Yes	Yes	No
Paradox Free	No	No	No	?	?	Yes	Yes	?

Legend:

Mc87	McCarthy: Generality in AI [38]
Mc93	McCarthy: Notes on formalizing context [40]
Gu91	Guha: A formalization of contexts [24]
Ba86	Barwise: Constraints and conditional information [8]
Sh91	Shoham: Varieties of context [52]
BM93	Buvač and Mason: Propositional logic of context [18]
AS93	Attardi and Simi: A formalization of viewpoints [5]
Ours	This report

Table 6.1: Comparison of the previous approaches and our approach

natural language applications.

- *Extensions to temporal domain:* In the statement of our proposal, we have not dealt with temporal relations. In Guha's work [24], most of the examples are related to time. As a future work, 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 [59] and PROSIT [42].

Bibliography

[1] P. Aczel. *Non-Well-Founded Sets.* Number 14 in CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, CA, 1988.

[2] V. Akman and M. Surav. Sets in context and related issues. Technical Report BU-CEIS-94-10, Department of Computer Engineering and Information Science, Bilkent University, Ankara, 1994. Available via anonymous ftp from `ftp.cs.bilkent.edu.tr` as `/pub/tech-reports/1994/BU-CEIS-9410.ps.z`.

[3] 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.

[4] P. A. Angeles. *Dictionary of Philosophy.* Harper and Row Publishers, New York, NY, 1981.

[5] G. Attardi and M. Simi. A formalization of viewpoints. Technical Report TR-93-062, International Computer Science Institute, University of California, Berkeley, CA, 1993.

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

[7] B. Partee, editor. *Montague Grammar.* Academic Press, New York, NY, 1976.

[8] 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. Reprinted in [10].

[9] J. Barwise. On the circumstantial relation between meaning and content. *Versus*, 44:23–39, 1986. Reprinted in [10].

[10] J. Barwise. *The Situation in Logic.* Number 17 in CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, CA, 1989.

[11] J. Barwise. Constraints, channels, and the flow of information. In P. Aczel, D. Israel, Y. Katagiri, and S. Peters, editors, *Situation Theory and Its Applications, III*, number 27 in CSLI Lecture Notes. Center for Study of Language and Information, Stanford, CA, 1993.

[12] J. Barwise and J. Etchemendy. *The Liar: An Essay on Truth and Circularity.* Oxford University Press, New York, NY, 1987.

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

[14] J. Barwise and J. Seligman. The rights and wrongs of natural regularity. In J. Tomberlin, editor, *Philosophical Perspectives 8: Logic and Language.* Ridgeview Press, Atascadero, CA, 1992. Available via anonymous ftp from `phil.indiana.edu` as `/pub/BarSel92.tex`.

[15] R. J. Brachman and H. J. Levesque, editors. *Readings in Knowledge Representation.* Morgan Kaufmann, Los Altos, CA, 1985.

[16] J. S. Bruner. On perceptual readiness. *Psychol. Rev.*, 64:123–152, 1957.

[17] S. Buvač, V. Buvač, and I. A. Mason. Metamathematics of contexts. *Fundamenta Informaticae*, 1994. To appear. Available via anonymous ftp from `sail.stanford.edu` as `/pub/buvac/context.ps`.

[18] 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.

[19] *Collins Cobuild English Language Dictionary.* William Collins and Sons, London, UK, 1987.

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

[21] F. Dretske. *Knowledge and the Flow of Information.* MIT Press, Cambridge, MA, 1981.

[22] M. Ersan, E. Ersan, and V. Akman. Commonsense aspects of buying and selling. Technical Report BU-CEIS-94-07, Department of Computer Engineering and Information Science, Bilkent University, Ankara, 1994. Available via anonymous ftp from `ftp.cs.bilkent.edu.tr` as `/pub/tech-reports/1994/BU-CEIS-9407.ps.z`.

[23] J. M. Gawron and S. Peters. *Anaphora and Quantification in Situation Semantics.* Number 19 in CSLI Lecture Notes. Center for the Language and Information, Stanford, CA, 1990.

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

[25] R. V. Guha. Context dependence of representations in CYC. Technical Report CYC 066-93, Microelectronics and Computer Technology Corporation, Austin, TX, 1993.

[26] R. V. Guha and D. B. Lenat. CYC: A midterm report. *AI Magazine*, 11:32–59, 1990.

[27] R. V. Guha and D. B. Lenat. Enabling agents to work together. *Communications of the ACM*, 37(7):127–142, 1994.

[28] P. J. Hayes. The second naive physics manifesto. In J. R. Hobbs and R. C. Moore, editors, *Formal Theories of the Commonsense World*, pages 1–36. Ablex, Norwood, NJ, 1985. Reprinted in [15].

[29] J. Hintikka. *Knowledge and Belief.* Cornell University Press, Ithaca, NY, 1962.

[30] G. Hirst. *Anaphora in Natural Language Understanding: A Survey.* Number 119 in Lecture Notes in Computer Science. Springer-Verlag, Berlin, Germany, 1981.

[31] A. S. Hornby. *Oxford Advanced Learner's Dictionary of Current English.* Oxford University Press, Oxford, UK, 1974.

[32] D. B. Lenat, R. V. Guha, D. Pratt, K. Pittman, W. Pratt, and K. Goolsbey. The world according to CYC, Part 4. Technical Report ACT-CYC-022-91, Microelectronics and Computer Technology Corporation, Austin, TX, 1991.

[33] D. Lewis. General semantics. In D. Davidson and G. Harman, editors, *Semantics of Natural Language*, pages 169–218. D. Reidel, Dordrecht, Holland, 1972. Reprinted in [7].

[34] V. Lifschitz, editor. *Formalizing Common Sense: Papers by John McCarthy.* Ablex, Norwood, NJ, 1990.

[35] J. Martins. Belief revision. In S. C. Shapiro, editor, *The Encyclopedia of Artificial Intelligence, volume 1*, pages 58–62. John Wiley, Chichester, UK, 1987.

[36] J. McCarthy. Programs with common sense. In *Proceedings of Teddington Conference on Mechanization of Thought Processes*, pages 75–91. Her Majesty's Stationary Office, London, UK, 1959. Reprinted in [15].

[37] J. McCarthy. Circumscription—a form of non-monotonic reasoning. *Artificial Intelligence*, 13:27–39, 1980. Reprinted in [34].

[38] J. McCarthy. Generality in artificial intelligence. *Communications of the ACM*, 30:1030–1035, 1987. Reprinted in [34].

[39] J. McCarthy. Artificial intelligence, logic and formalizing common sense. In R. Thomason, editor, *Philosophical Logic and Artificial Intelligence*, pages 161–190. Kluwer Academic Press, Dordrecht, Holland, 1989.

[40] J. McCarthy. Notes on formalizing context. In *Proceedings of the Thirteenth International Joint Conference in Artificial Intelligence*, Chambery, France, 1993. Available via anonymous ftp from `sail.stanford.edu` as `/pub/jmc/context-2.tex`.

[41] R. Montague. Syntactical treatment of modalities, with corollaries on reflexion principles and finite axiomatizability. *Acta Philosophica Fennica*, 16:153–167, 1963.

[42] 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.

[43] W. V. O. Quine. Propositional objects. In *Ontological Relativity and Other Essays.* Columbia University Press, New York, NY, 1969.

[44] R. Thomason, editor. *Formal Philosophy: Selected Papers of Richard Montague.* Yale University Press, New Heaven, CT, 1974.

[45] R. Reiter. On closed world data bases. In H. Gallaire and J. Minker, editors, *Logic and Databases.* Plenum Press, New York, NY, 1978.

[46] R. Reiter. On reasoning by default. In *Proceedings of Theoretical Issues in Natural Language Processing-2 (TINLAP-2)*, pages 210–218, University of Illinois at Urbana-Champaign, IL, 1978. Reprinted in [15].

[47] E. Rosch. Principles of categorization. In E. Rosch and B. B. Lloyd, editors, *Cognition and Categorization*, chapter 2. Lawrence Erlbaum Associates, Hillsdale, NJ, 1978.

[48] J. Seligman. Perspectives in situation theory. In R. Cooper, K. Mukkai, and J. Perry, editors, *Situation Theory and Its Applications, I*, number 22 in CLSI Lecture Notes, pages 147–191. Center for Study of Language and Information, Stanford, CA, 1990.

[49] J. Seligman. Structure of regularity. Drafts of Lecture Notes, Indiana University Logic Group, Bloomington, IN, 1993. Available via anonymous ftp from `phil.indiana.edu` as `/pub/SORnotes/*`.

[50] P. Sells and T. Wasow. Anaphora. Privately obtained from the second author, 1993.

[51] Y. Shoham. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence.* MIT Press, Cambridge, MA, 1988.

[52] 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.

[53] E. Shortliffe. *MYCIN: Computer-Based Medical Consultations.* Elsevier, New York, NY, 1976.

[54] R. Stalnaker. *Inquiry.* MIT Press, Cambridge, MA, 1984.

[55] M. Surav and V. Akman. Axiomatic set theories and common sense. Technical Report BU-CEIS-94-11, Bilkent University, Ankara, 1994. Available via anonymous ftp from `ftp.cs.bilkent.edu.tr` as `/pub/tech-reports/1994/BU-CEIS-9411.ps.z`.

[56] *The Merriam-Webster Dictionary.* Pocket Books, New York, NY, 1974.

[57] E. Tın. Computing with causal theories. Master's thesis, Department of Computer Engineering and Information Science, Bilkent University, Ankara, 1990.

[58] E. Tın and V. Akman. Computing with causal theories. *International Journal of Pattern Recognition and Artificial Intelligence*, 6:699–730, 1992.

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

[60] E. Tın and V. Akman. Situated processing of pronominal anaphora. In *2. Konferenz Verarbeitung natürlicher Sprache (KONVENS 94)*, Vienna, Austria, 1994. To appear.

[61] R. W. Weyhrauch. Prolegomena to a theory of mechanized formal reasoning. *Artificial Intelligence*, 13:133–170, 1980. Reprinted in [15].