

Theories and uses of context in knowledge representation and reasoning

Paolo Bouquet^a Chiara Ghidini^b
Fausto Giunchiglia^{a,c} Enrico Blanzieri^d

^a*Department of Information and Communication Technology University of Trento*

^b*Department of Computer Science, University of Liverpool*

^c*ITC-irst, Trento*

^d*Department of Psychology, University of Torino*

bouquet@cs.unitn.it chiara@csc.liv.ac.uk
fausto@dit.unitn.it blanzier@dit.unitn.it

Abstract

This paper discusses the uses of context in knowledge representation and reasoning (KRR). We propose to partition the theories of context brought forward in KRR into two main classes, which we call *divide-and-conquer* and *compose-and-conquer*. We argue that this partition provides a possible explanation of why in KRR context is used to solve different types of problems, or to address the same problems from very different perspectives. The problems we use to illustrate this point are the problem of generality, the formalization of propositional attitudes, and knowledge and data integration.

Key words:

knowledge representation, contextual reasoning, theory of context, uses of context

Abbreviated title:

context in KR and reasoning

short bio:

Paolo Bouquet has a degree in Philosophy from the University of Milan, and a PhD. in Philosophy of Science from the University of Genoa. Currently, he is assistant professor in computer science at the University of Trento and collaborates with the Cognitive Science Laboratory at Rovereto (Trento). He

is Steering Committee Chair of the International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT). His research is at the intersection between artificial intelligence, logic, epistemology, philosophy of language, and cognitive science. His recent publications concern formal theories of contextual reasoning, context-aware applications, knowledge management, formal semantics for indexical languages.

Chiara Ghidini obtained a PhD in Computer Science in 1998 in a joint programme between the University of Rome "La Sapienza" and the University of Trento, Italy. She is currently a lecturer at the University of Liverpool, UK. Since finishing her PhD, Dr. Ghidini has rapidly established herself as a significant member of the Context community. Her main research interests have been in logic and knowledge representation. She has contributed to research in formal theories of contextual reasoning and their applicability to multi-agent and distributed information systems that has led to several publications in leading journals and conferences.

Fausto Giunchiglia is Professor of Computer Science at the University of Trento, Department of Information and Communication Technology and Senior Consultant at ITC-IRST, Automated Reasoning Systems Division of Trento. He is well-known for his research in several areas of Computer Science. He participates to the IJCAI Board of Trustees; he is Member of the Editorial Board of the Book series "Frontiers in AI and its applications", IOS Press; he is Editorial Board Member of AICOM - AI Communications, IOS Press. He has been Advisory Board Member and Associate Editor of the Journal of Artificial Intelligence Research (JAIR). He has been co-organizer of many conferences (e.g. IJCAI, KR&R, FLOC).

Enrico Blanzieri has a degree in Electronic Engineering from the University of Bologna, Italy and a Ph.D in Cognitive Science from Centro di Scienza Cognitiva, Università e Politecnico di Torino. From 1997 to 2000 he was within SRA division at ITC-irst Trento. He has been researcher at the University of Torino and he is currently researcher in computer science at the University of Trento. His main areas of research are Instance-Based Learning and cognitive modelling of communicative and reasoning phenomena. His recent work is in the intersection between machine learning, multi-agent systems and representation of shared knowledge.

1 Introduction

The notion of context plays a crucial role in such different disciplines as pragmatics, natural language semantics, linguistics, cognitive psychology, and artificial intelligence (AI). Studies of context have generated a highly interdisciplinary field and the proceedings of the international and interdisciplinary conference on “Modeling and Using Context” offer a good illustration of this claim (Akman et al., 2001; Bouquet et al., 1999). Even within AI, context is used in many different areas. In natural language processing, context is used to assign an interpretation to assertions and resolve ambiguities. In information retrieval, context helps to refine the queries made by users. In distributed AI, context is used as a flexible formal tool for the design of systems of autonomous agents. In human-machine interaction, context is used to design context-sensitive applications and interfaces. In this paper, we focus on the use of context in an area of AI called *knowledge representation and reasoning* (KRR), whose aim is to devise languages for representing what (intelligent) programs or agents know about their environment, and for representing the reasoning processes that allow them to derive new knowledge from what they already know.

In KRR, a notion very similar to context, called *LSpair*, was first introduced by Weyhrauch in his *Prolegomena to a theory of mechanized formal reasoning* (1980). His goal was to implement the epistemological part of McCarthy’s *Advice Taker*, a program designed to possess abilities that in human beings would be called common sense. A fundamental assumption of the Advice Taker project was that formal logic was an appropriate tool for modeling and studying the properties of such a program. In particular, McCarthy held that the program’s knowledge was to be represented as a logical theory, and that reasoning was to be modeled as inference in such a theory. Weyhrauch introduced *LSpairs* (that he later called “contexts”) as crucial devices for the mechanization of these ideas. A context was thought of as a finite representation of a logical theory suitable for being mechanized. Context was viewed as the building block of a theory of mechanized reasoning.

It was only in the late 80’s/early 90’s, however, that context became a widely discussed issue. Independently from each other, and with different motivations, Giunchiglia (1993) and McCarthy (1987; 1993) started to work on a formal theory of context, whose goal was to explain the properties of context and contextual reasoning in a systematic way.

Since then, context has been used in different types of applications in KRR. CYC, the largest common sense knowledge base ever built, implements and exploits an explicit notion of context (Lenat et al., 1990; Guha, 1991). Context also plays an important role in the formalization of reasoning about beliefs (Giunchiglia et al., 1993; Giunchiglia and Giunchiglia, 1996; Benerecetti et al., 1998a; Ghidini, 1999), metareasoning and propositional attitudes (Giunchiglia and Serafini, 1994), reasoning with viewpoints (Attardi and Simi, 1995), and

reasoning about action (Bouquet and Giunchiglia, 1995). Other important uses are in the modeling of agents and multi-agent systems (Benerecetti et al., 1998b; Cimatti and Serafini, 1995), dialog, argumentation, (Parsons et al., 1998; Noriega and Sierra, 1996), and in the integration of heterogeneous and autonomous knowledge and databases (Farquhar et al., 1995; Mylopoulos and Motschnig-Pitrik, 1995; Ghidini and Serafini, 1998c,b).

The goal of the paper is twofold: to analyze the theories of context that have been proposed in KRR, and to apply our findings to describing the different uses of context in KRR. In the first part, we show that the theories of context proposed in KRR can be divided into two general types: the first, which we call *divide-and-conquer*, sees context as a way of partitioning a global model of the world into smaller and simpler pieces; the second, which we call *compose-and-conquer*, sees context as a local theory of the world in a network of relations with other local theories. We briefly discuss the relationships between the two types of theory and argue that the second is more general than the first. In the second part of the paper, we show that each type of theory leads quite naturally to different uses of context, namely that some problems are more naturally addressed by a divide-and-conquer theory and others by a compose-and-conquer theory. This is illustrated by discussing examples like the problem of generality, the formalization of propositional attitudes, and knowledge and data integration. Readers interested to more complete surveys may refer to Akman and Surav (1996) for formal models of context, and to Brézillon (1999) for applications of context.

2 Two types of theory of context in KRR

The goal of KRR is to provide and study formal languages that can be used to represent what an agent of a certain kind knows about the world, and to show how this knowledge can be used in a reasoning process to infer new knowledge from that already available. With respect to this goal, many researchers believe that a completely general representation of knowledge is impossible in practice, and – more interestingly – perhaps not even desirable. Indeed, whatever language and facts we choose to represent, there is always a situation in which the stated facts or the language itself are not adequate. Here’s how McCarthy expressed this intuition in his well-known paper on generality in AI (1987:1032):

“Whenever we write an axiom, a critic can say that the axiom is true only in a certain context. With a little ingenuity the critic can usually devise a more general context in which the precise form of the axiom doesn’t hold. Looking at human reasoning as reflected in language emphasizes this point. Consider axiomatizing *on* so as to draw appropriate consequences from the information expressed in the sentence, ‘The book is on the table’. The critic may propose to haggle about the precise meaning of *on*, inventing difficulties

about what can be between the book and the table, or about how much gravity there has to be in a spacecraft in order to use the word *on* and whether centrifugal force counts. Thus we encounter Socratic puzzles over what the concept means in complete generality and encounter examples that never arise in life. There simply isn't a most general context.

Conversely, if we axiomatize at a fairly high level of generality, the axioms are often longer than is convenient in special situations. Thus humans find it useful to say, 'The book is on the table', omitting reference to time and precise identification of what book and what table. [...]

A possible way out involves formalizing the notion of context [...]"

The quotation makes clear that an important motivation for studying formal theories of context is the need for problem-tailored representations. Indeed, this should allow a reasoning system to avoid the use of unnecessarily complicated theories, and at the same time provide the ability to "jump" to a more general representation if the one in use proves to be inadequate.

The idea of using context as a tool to "localize" reasoning to a subset of the facts known by an agent (ideally, to the "right" set of facts in a given circumstance) is also one of the motivations brought forward by Giunchiglia in one of his first papers on context (1993:345):

"It is widely agreed on that most cognitive processes are *contextual* in the sense that they depend on a set of variables which constitute the environment (or context) inside which they are carried on [...] Our basic intuition is that reasoning is usually performed on a subset of the global knowledge base; we never consider *all we know* but only a very small subset of it [...]

We take a *context c* to be that subset of the complete state of an individual that is used for reasoning about a given goal."

A third illustration of this widely shared view on context is Lenat's (1999:3) account of why context was introduced into the CYC common sense knowledge base (see Section 3.3.3):

"During the 1984-1989 time period, as the Cyc common sense knowledge base grew ever larger, it became increasingly difficult to shoehorn every fact and rule into the same flat world. Finally, in 1989, as Cyc exceeded 100,000 rules in size, we found it necessary to introduce an explicit context mechanism. That is, we divided the KB up into a lattice of hundreds of contexts, placing each Cyc assertion in whichever context(s) it belonged."

A further illustration of this idea is Dinsmore's book on partitioned representation. He sees partitioned representations as the functional counterpart of the notion of "mental space" as defined by Fauconnier (1985), and defines them as context-dependent representations of the world (Dinsmore 1991:45):

"Functionally, partitioned representations illustrate the principle of *divide and conquer* in mental representation. Rather than a large homogeneous representation or set of representations with unmanageable possibilities for synthesis in reasoning, partitioned representations make use of many iso-

lated context-dependent representations with locally circumscribed [...] opportunities.”

Interestingly enough, KRR seems to share this intuition with other related areas. Two examples will illustrate this “family resemblance”. Sperber and Wilson, in their book on relevance (1986:15), express a similar intuition from a psycholinguistic perspective:

“The set of premises used in interpreting an utterance [...] constitutes what is generally known as the context. A context is a psychological construct, a subset of the hearer’s assumptions about the world.”

And Kokinov, in his paper on a dynamic approach to context modeling, says that (Kokinov 1995:200):

“[c]ontext is the set of *all* entities that influence human (or system’s) cognitive behavior on a particular occasion.”

Despite the evidence of a shared intuition, we argue that there are at least two different types of theories of context that have been proposed in KRR:

- the first sees a context as a way of partitioning (and giving a more articulated internal structure to) a global theory of the world;
- the second sees a context as a local theory, namely a (partial, approximate) representation of the world, in a network of relations with other local theories.

According to the first view, which we call *divide-and-conquer*, there is something like a global theory of the world. This global theory has an internal structure, and this structure is articulated into a collection of contexts. According to the second view, which we call *compose-and-conquer*, there is not such a thing as a global theory of the world, but only many local theories. Each local theory represents a viewpoint on the world. Also, there may exist relations between local theories that allow a reasoner to (partially) compose them into a more comprehensive view. In the rest of this section, we analyze these two views of context, and present in some detail two formalizations of context as a practical illustration of the two types of theory: the Propositional Logic of Context (LoC) (Buvač and Mason, 1993), which illustrates the divide-and-conquer type, and Local Models Semantics / MultiContext Systems (LMS/MCS) (Ghidini and Giunchiglia, 2001; Giunchiglia and Serafini, 1994), which illustrates the compose-and-conquer type. We will briefly review also Dinsmore’s theory of partitioned representations, for it provides an interesting variation of a divide-and-conquer theory. Of course, there are several other (formal) theories of context in KRR. Examples are structured contexts with fibred semantics (Gabbay and Nossum, 1997), and the type theoretic foundation for context (Thomason, 1999). However, we decided to focus on LoC and LMS/MCS for two reasons: first, they provide the clearest illustrations of the two types of theories we described above, and second, of all the existing theories of context, they are the most extensively used in KRR.

Partitioning a global theory of the world can have two different meanings: (i) that the collection of facts globally available is partitioned into smaller subsets, each of which describes knowledge about some domain, or knowledge that is needed to solve a specific problem, or (ii) that the same set of facts can be given different descriptions, each at a different level of detail, depending on what is implicitly assumed. For example, the fact that at a time T_0 a block A is on a block B can be represented as $on(A, B, T_0)$, but also as $on(A, B)$ in a context in which it is implicitly assumed that the time is T_0 .

In general, a divide-and-conquer theory has the following form:

- given a (global) representation language L , the facts that are true in a given context c can be isolated (“localized”) and treated as a distinct collection of facts with respect to the facts belonging to different contexts;
- there are hierarchical relations between contexts that allow reasoning to “climb” from a context to a more general context in which the dependence of a fact on a context is explicitly stated (and possibly reasoned about);
- finally, there are lateral (i.e. non-hierarchical) relations between facts of different contexts (for example, one would like to be able to represent the relation between the fact $on(A, B, T_0)$ in the context bw of the blocks’ world and the formula $on(A, B)$ in a context $specializes(bw, T_0)$ that specializes the context of the blocks world to time T_0).

[INSERT FIGURE 1 HERE]

LoC, as originally described by McCarthy and then formalized by Buvač and Mason (1993), is a possible way of capturing the general intuitions of a divide-and-conquer theory. The building blocks of LoC can be described as follows (see Figure 1):

- first of all, any formula can only be asserted in some context (namely, there are no context independent formulae). The fact that a formula p is asserted in a context c is written as $c : p$. Context sequences are used to represent nested contexts. Context sequences allows distinctions, say, between the context of car racing in the context of the 50’s and the context of car racing in the context of the 90’s. Notationally, context sequences are represented by a sequence $c_1 \dots c_n$ of context names, or by \bar{c} .
- contexts are reified as first-class objects, which means that they are objects about which we can make assertions in the language of the theory. The representation language L is hence enriched with a set of context names c_1, c_2, \dots ;
- the most important statements about a context c are made through the

formula $ist(c, p)$. Intuitively, it means that the formula $p \in L$ is true in the context c . Buvač and Mason (1993) and Buvač (1996) treat ist as a modality.

- the main hierarchical relation between facts belonging to different contexts is that between the fact p stated in a context c , and the formula $ist(c, p)$ stated in some outer context c' . The relation is the following:
 - if $ist(c, p)$ can be proved in c' , then we can always “enter” the context c and assert $c'c : p$. If c is the context of car racing, this allows us to “enter” c and restrict reasoning only to the facts that are true in it;
 - if p can be proved in c , then we can always “leave” (or *transcend*) the context c and make the explicit assertion $ist(c, p)$ in the outer context c' (i.e. $c' : ist(c, p)$).

These relations define the nested structure of Figure 1. Indeed, leaving (or transcending) a context is tantamount as moving from a box to the box immediately outside; whereas entering a context is the move from outside to inside;

- other (lateral) relations between facts belonging to different contexts are stated through the so-called *lifting axioms* (see Guha (1991) for this notion). The general form of a lifting axiom is

$$ist(c, p) \iff ist(c', p')$$

In Figure 1, lifting axioms are represented as dotted lines connecting boxes. One can use a lifting axiom to say that, for example, in the context of the Sherlock Holmes stories, it is true that Holmes lived near Victoria Station if and only if, in the context of the actual city of London, Victoria Station is near Baker Street;

- finally, there is no such a thing as an outermost context. This reflects McCarthy’s intuition that we can never resolve all contextual dependencies of a fact. More technically, this means that we can always transcend a context c and move to a more general context in which facts about c can be asserted (including those making some of its implicit assumptions explicit). In Figure 1, this is represented by the external dotted boxes.

These intuitions are formalized in Buvač and Mason’s propositional logic of context (1993), followed by Buvač’s first order formalization of context (1996). For the sake of simplicity, we only describe the propositional part, called LoC. In LoC, we start with a propositional language L (which includes a collection of context names and the modality ist). Roughly speaking, a model \mathfrak{M} for LoC associates a set of partial truth assignments to each sequence of contexts \bar{c} (possibly of length one). Satisfiability is defined with respect to \bar{c} . The idea is that partial truth assignments capture the fact that in different contexts there are different sets of meaningful formulae. Indeed, starting from a unique language L , a model \mathfrak{M} defines a function, called *vocabulary*, that associates to each context c a subset of L , which is the set of meaningful formulae in c . Obviously, satisfiability and validity of formulae are defined only for these models

that provide enough vocabulary, namely the vocabulary which is necessary to evaluate a formula in a context.

[INSERT FIGURE 2 HERE]

Buvač and Mason propose an Hilbert style axiomatization of validity for the logic of context, which is presented in Figure 2. (PL) says that all propositional tautologies are valid in every context \bar{c} . Axiom (K) imposes that predicate *ist* satisfies properties analogous to those of the modality \Box in a modal system K , while axiom (Δ) forces the truth of *ist* formulae (i.e. formulae of the form $ist(c, p)$) to be independent of the assignments of the contexts in which they occur (for a short discussion of the axiom (Δ), see Bouquet and Serafini (2001)). (MP) is the usual rule of Modus Ponens.

The rule (CS) is very important. It formalizes the hierarchical relationship between contexts in LoC. This relationship, which is part of the logic itself, is the mechanism that allows transcendence of a context. As we already said, this corresponds to moving one step outward in Figure 1.

Very briefly, we present also Dinsmore's theory of partitioned representations (PR) (1991), a logic strictly related to Fauconnier's work on *mental spaces* (1985). As we will show, PR can be viewed as another instance of a divide-and-conquer theory. Despite some terminological differences, LoC and PR share most of their structure. However, there is an important difference: while in LoC the process of transcending a context is open-ended and there is no outermost context, in PR we eventually reach a special space, called BASE, which cannot be further transcended. BASE can therefore be considered as a sort of outermost context.

In PR, a sentence is always asserted in a space. A space represents some logically coherent situation or potential reality, where various propositions are treated as true, objects are assumed to exist, and relations between objects are supposed to hold (e.g. belief spaces, hope and wish spaces, fictional, dream, and pretense spaces, spaces representing specific places, times and situations, spaces representing the scope of certain existential assumptions, spaces expressing generalizations, and spaces representing the implications of certain propositional assumptions, either conditional or counterfactual). In Figure 3, spaces are represented as rectangles.

[INSERT FIGURE 3 HERE]

Each space has exactly one primary context. A *primary context* is defined as a

function that maps the truth of a proposition in one space onto the satisfaction of a (more complex) proposition in another space. Suppose, for example, that the sentence “Mr. Bush is the President of the US” is asserted in a space named S_1 and that “Warren believes that [S_1]” is the primary context of S_1 (here we are using a notation which is slightly different from the original). This allows us to map the truth of “Mr. Bush is the President of the US” onto the truth of the (more complex) sentence “Warren believes that [Mr. Bush is the President of the US]”, which, in turn, is asserted in some other space. Of course, the semantics of “Mr. Bush is the President of the US” would be very different in a space S_2 whose primary context was something like “In the Sherlock Holmes stories [S_2]”. This process is called *context climbing*, and corresponds very closely to the idea of transcending (leaving) a context in LoC.

Via context climbing, PR allows us to reach a special space, called BASE, which functions as an outermost space. BASE is the only space that does not have a primary context. Even though Dinsmore dislikes this interpretation, formally speaking, BASE represents a “de-contextualized” representation of the world. Indeed, it is only in BASE that assertions are given an interpretation. As a consequence, if BASE is not reachable (via context climbing) from a space S , then the assertions of that space are left without a truth value.

Dinsmore introduces also a notion of *secondary context*, which allows for lateral mappings. Intuitively, a mapping is a consequence of the semantics of the primary contexts involved. In other words, a secondary context opens a channel of communication between two spaces. For instance, if S_1 models Warren’s beliefs about Bush and S_2 Warren’s beliefs in general (no matter about what), then we can imagine that the facts asserted in S_1 will be inherited by S_2 . Inheritance is just one – perhaps the simplest – use of secondary context in PR.

Although LoC is formally richer than PR, it is easy to see that there are many similarities between them. They share the general structure of a compose-and-conquer theory of context: a way of localizing collections of facts, a hierarchical relation between contexts built into the logic, and a way of defining non-hierarchical (lateral) mappings between facts belonging to different contexts. We presented both theories because they offer different solutions to the problem of de-contextualization: open-ended in LoC and bounded in PR.

2.2 “Compose-and-conquer” theories of context

Compose-and-conquer theories start from the assumptions that local (domain specific, goal directed) theories of the world are the building blocks of what an agent knows, and that the totality of the agent’s knowledge is given by composing such local theories through a collection of rules that connect them into a more comprehensive (but still partial) representation of the world.

The general structure of a compose-and-conquer theory can be described as in Figure 4. First of all, each box is a local theory. A local theory is not a partition of a bigger (global) theory, but a full-blown theory which represents knowledge about some portion of the world (partiality), at some level of detail (approximation), from a given perspective. Examples are: domain theories (e.g. about air travel, cars, sports, cooking), snapshots of a dynamic situation (e.g. the state of a chess game, the current situation during the execution of a plan), representation of a physically limited portion of the world (e.g. the location of physical objects in a room, the location of restaurants in New York), and beliefs ascribed to another agent or group of agents (e.g. the beliefs that John ascribes to Mary, John’s beliefs about the beliefs that Mary ascribes to him).

[INSERT FIGURE 4 HERE]

Second, in a compose-and-conquer theory, there are no a priori relations between contexts. This is a major difference from divide-and-conquer theories. As we said, hierarchical relations are “hardwired” in systems like LoC and PR, and the reason for this is that contexts are viewed as chunks of a bigger (global) theory of the world; in a sense, the global model is a sort of road map that says how contexts are related to each other. In compose-and-conquer theories, on the other hand there is no predefined road map, and contexts are autonomous theories, though partial and approximate. Of course, this does not mean that there are no relations between contexts, but only that these relations are established on a peer-to-peer basis, as a collection of constraints on what can (or cannot) be true in a context given that there is some relation with what holds in another context. A special case of a peer-to-peer relation is a hierarchical relation (e.g. transcendence), but it is interpreted as a constraint between two autonomous local theories. For example, there may be a constraint between the truth of a fact p in a context c and the truth of a fact $ist(c, p)$ in a context c' , but this is interpreted as the fact that any local model of c that satisfies p is incompatible with any local model of c' that does not satisfy $ist(c, p)$, and vice versa. In this respect, hierarchical relations in compose-and-conquer theories are assimilated with all the other relations, and so the logic can express hierarchies of contexts without the need to predefine them in the logic itself.

A clear formalization of a compose-and-conquer theory of context is Ghidini and Giunchiglia’s *Local Models Semantics* (LMS) (2001), together with its proof-theoretical counterpart, namely, Giunchiglia and Serafini’s *MultiContext Systems* (MCS) (1994). LMS is based on two very general principles, that we restate as follows:

- *principle of locality*: reasoning always happens in a local theory (a context);
- *principle of compatibility*: there may be compatibility constraints between

the reasoning processes that happen in different contexts.

Ghidini and Giunchiglia (2001) provide both a model-theoretic (LMS) and a proof-theoretic (MCS) formalization of these two principles (even though the first proof-theoretic version was originally provided by Giunchiglia and Serafini (1994)). In LMS, one starts with a family of languages L_1, \dots, L_n, \dots , where each L_i is the representation language of a context c_i . Each language L_i has its set of models M_i . Every subset M_{T_i} of M_i satisfies a set of formulae, each corresponding to a different choice of the theory (set of true facts) T_i associated with c_i . By abuse of notation, we will use the symbol c_i (possibly with different subscripts), to mean either the theory associated with context c_i or a context embedded in a structure of relationships with other contexts. Once the theory T_i associated with c_i is fixed, a model belonging to M_{T_i} is called a *local model* of c_i . Going back to Figure 4, this means that each context (box) c_i depicted in that figure is formalized by the set M_{T_i} of models of L_i which satisfies the axioms in c_i . Relations between two contexts (dotted lines between boxes in Figure 4) are represented by compatibility constraints, which state that the truth of a formula Φ in c_1 is related to the truth of the formula Ψ in c_2 (the case of multiple contexts and multiple compatibility constraints is a straightforward generalization). This is achieved by imposing that sets of local models \mathbf{c}_1 and \mathbf{c}_2 of the two contexts c_1 and c_2 are such that

$$\text{if } \mathbf{c}_1 \text{ satisfies } \Phi, \text{ then } \mathbf{c}_2 \text{ satisfies } \Psi \quad (1)$$

where the notion of satisfiability of a formula in a set of local models is the same as the notion of satisfiability of a formula in the theory associated to c_i . Pairs $\langle \mathbf{c}_1, \mathbf{c}_2 \rangle$ satisfying Equation (1) are said to belong to a compatibility relation and define a model for the pair of contexts $\{c_1, c_2\}$.

The proof-theoretic counterpart of LMS is called *MultiContext Systems* (MCS). A MCS is a pair $MC = \langle \{c_i\}, BR \rangle$, where $\{c_i\}$ is a set of axiomatic formal theories (namely triples of the form $c_i = \langle L_i, \Omega_i, \Delta_i \rangle$), and BR is a set of *bridge rules*. Bridge rules are rules whose premises and conclusion belong to different contexts. For instance, the bridge rule corresponding to the compatibility constraint described above would be the following:

$$\frac{c_1 : \Phi}{c_2 : \Psi}$$

where $c_1 : \Phi$ is the premiss of the rule and $c_2 : \Psi$ is the conclusion. Obviously, bridge rules are conceptually different from *local rules* (i.e., rules in Δ_i). The latter can be applied only to formulae of L_i , whereas the former have the premises and the conclusion that belong to different contexts. Intuitively, bridge rules allow for the MCS version of compatible derivations. A deduction in a MCS is a tree of local deductions, obtained by applying only rules in Δ_i , concatenated with one or more applications of bridge rules (see Giunchiglia and Serafini (1994) for a technical treatment).

Using the machinery of compatibility in LMS (or bridge rules in MCS), a wide range of relations between contexts can be formalized. Some examples are:

- suppose a context c represents John’s beliefs at time t , and that it contains the sentence “I’m hungry”. If c_1 represents John’s beliefs at time $t + 1$ (e.g. the day after), then there exists a relationship between the two contexts such that the sentence “Yesterday I was hungry” must be true in c_1 ;
- suppose c_2 represents Mary’s beliefs at time t , and that John tells Mary “I’m hungry”. Then the relationship between c and c_2 is such that the sentence “He is hungry” (or “John is hungry”, if Mary knows that the speaker’s name is John) must be true in c_2 ;
- suppose c_3 represents the positions of the object in a room (including John and Mary) at time t and that the sentence “John is near Mary” is true in c_3 . If c_4 represents the location of the same objects at different times, then the sentence “At time t , John was near Mary” must be true in c_4 .

All these relations have the following form: if a sentence p is true in a context c , and c is in a given relation R with c' , then the sentence p' must be true in c' . Of course, a special case is when there is no relation between the two contexts. In this case there are no constraints on what is true in the two contexts.

It is easy to see why LMS/MCS is a “compose-and-conquer” logical framework. Model-theoretically, the idea is that contexts can be composed through compatibility relations, which allow the exclusion of all local models that are not compatible with the known relationships between two local theories. Proof-theoretically, the idea is that bridge rules allow us to derive in a context more facts than would be derivable if the context was taken in isolation. The reason is that these further facts are derivable precisely because of the relations that contexts have with each other. However, both model- and proof-theoretically, there is no assumption that there is a global model of the world.

To sum up, there are very significant conceptual and formal differences between divide-and-conquer and compose-and-conquer theories. Indeed, in a compose-and-conquer theory:

- there is not such a thing as a general representation language L . The representation language is context-dependent, as it reflects the “ontology” implicitly assumed in a local representation;
- denotation and truth are by definition contextual, as they are defined with respect to the language and models of each context;
- reasoning is local by definition, as it always happens in a context. Since there is no single logical space that contains (not even potentially) all that an agent knows, reasoning can only happen in the small logical spaces that correspond to a context. This means that different contexts may even correspond to different reasoning rules;
- relationships between different contexts are not necessarily stated and used at a meta-level. They can be viewed as constraints on what can be locally derived in a context (see below for technical details about this).

Despite the differences above, Bouquet and Serafini (2001) prove that LMS and MCS can be used to “simulate” a divide-and-conquer approach. In particular, it is shown that LMS/MCS can subsume LoC. Intuitively, a divide-and-conquer approach uses context as a mechanism for partitioning a global representation into logical spaces that are smaller and simpler than the global knowledge space of a program (or, more in general, of an agent). This can be easily done in a logic where each context is described by its own local language and semantics. The operations of entering and leaving a context can be modeled as a specific compatibility relation imposing that p holds in the context c if (and only if) $ist(c, p)$ holds in \bar{c} . Analogously, lifting axioms can be modeled as compatibility constraints between local theories. This intuitively justifies the claim that the principles of locality and compatibility are the most general principles of contextual reasoning, and that a logic based on these principles is general enough to provide a suitable basis for context-based KRR applications.

3 Uses of context in KRR

In the following, we focus on the use of context in specific areas of KRR. Our aim is to illustrate the idea that some problems are more naturally addressed from a divide-and-conquer perspective, others from a compose-and-conquer perspective. The examples we consider are: the problem of generality, the formalization of propositional attitudes, and knowledge and data integration.

3.1 *The problem of generality*

The so-called problem of generality (identified by McCarthy in his paper on *Generality in artificial intelligence* (1987)) and its related problem, the qualification problem, are a typical point of contact between context and KRR. Giunchiglia et al. (1996) elaborate a specific version of the problem of generality: the problem of dealing with expected and unexpected obstacles in the so-called Glasgow-London-Moscow (GLM) example. This problem was first proposed by McCarthy in the unpublished note *Overcoming an unexpected obstacle* (1991):

“You are planning a trip from Glasgow to Moscow via London. You would like to build the plan maybe without having to think of all the details, i.e. by working in a fairly approximated theory. For instance you are willing to consider the fact that you must have a ticket in order to get on a plane but not the fact that the flight could be canceled. However you want to be able to revise your plan if an expected obstacle arises (e.g. you do not have the ticket because you have lost it) and more particularly if an unexpected obstacle arises (e.g. the flight is cancelled).”

Here the problem is the trade-off between needed generality and an excess of generality. The requirement of building a plan “without having to think of all

the details” goes in the direction of finding the appropriate context containing only the information which is needed to solve the specific problem of planning a trip from Glasgow to Moscow via London. However, the fact that “you want to be able to revise your plan if an expected obstacle arises [...] and more particularly if an unexpected obstacle arises” requires the ability to change to a less general context, containing also the information which is needed to overcome the unexpected obstacle. Depending on the context, the “same” problem of planning a trip from Glasgow to Moscow via London can be given several different representations with different degree of generality. If, on the one hand, a more general representation can be applied to a larger class of circumstances, on the other hand, too much generality is a problem from the standpoint of implementing a reasoning system. In many contexts, some information can be left implicit that in other contexts must be included. Another prototypical example of the problem of generality is McCarthy’s (1993:557) ‘above–theory’ example:

“Consider a context *above-theory*, which expresses a static theory of the blocks world predicates *on* and *above*. In reasoning about the predicates themselves it is convenient not to make them depend on situations or on a time parameter. However, we need to *lift* the results of *above-theory* to outer contexts that do involve situations or times.”

As a consequence, the *above-theory* context contains very simple axioms on the blocks world of the form:

$$on(x, z) \supset above(x, z) \tag{2}$$

$$above(x, y) \wedge above(y, z) \supset above(x, z) \tag{3}$$

These axioms say that an object x is above an object z if either x is on z or x is above an object y which, in turn, is above z . Most times, these two axioms are sufficient for reasoning about the property of being above in the blocks world. However, there are cases where these two axioms are not general enough. For instance, (3) is true only if $above(x, y)$ and $above(y, z)$ are true at the same time. If we need to reason about the property of being above in a context c that does involve situations or times, McCarthy suggests the use of lifting rules to export axioms from the *above-theory* to c and to add a parameter for the time to the predicates *on* and *above*. Axioms (2) and (3) are then exported to c as:

$$on(x, z, t) \supset above(x, z, t) \tag{4}$$

$$above(x, y, t) \wedge above(y, z, t) \supset above(x, z, t) \tag{5}$$

Summarizing, McCarthy suggests solving the problem of generality with a formalization of context which allows us to use the “right” axioms in the “right” context, e.g. the less general axioms (2) and (3) if the context allows us to disregard the time, and the more general axioms (4) and (5) in a context where time is relevant.

The formalization of context proposed by McCarthy and his group, including Buvač and Mason's (1993) logic of context, is meant to deal particularly with the problem of generality. The reader interested in this approach may refer to the work of McCarthy (1993) and Guha (1991) for discussion of the general underlying intuitions and many motivating examples, and to the work of Buvač and Mason (1993) for the logical framework. From the description of the problem of generality given above, it is easy to see how the notion of context used to solve this problem is naturally addressed from a divide-and-conquer perspective. There is a global model of the world. However, too much knowledge is a problem from the standpoint of implementing reasoning or planning systems. Contexts provide the solution, as they are devices used to focus on (smaller, simpler) portions of such a model. Transcendence and lifting are the mechanisms used to relate contexts at different levels of generality.

3.2 Modeling propositional attitudes

Contexts have been extensively used to formalize mental states. It is not very surprising to discover that this problem has been mainly addressed from a compose-and-conquer perspective. In fact, contexts have been used in this area because they provide a tool for the representation of a (partial, approximate) theory from an individual's perspective. As a consequence, most of the work we describe here is formalized using Local Models Semantics and MultiContext systems.

3.2.1 VIEWGEN

Ballim, Wilks and colleagues proposed VIEWGEN as a framework for modeling agents which have a representation of beliefs, intentions, and goals of other agents involved in a dialogue (Ballim and Wilks, 1991; Lee and Wilks, 1996). VIEWGEN assumes that each agent taking part in a dialogue has a belief environment which includes attitudes about what other agents believe, want, and intend. Such attitudes are represented as a nested structure. Each nesting contains propositions which may be grouped by a particular topic or stereotype. The particular topic is given on the top left corner of the environment while the holder of a belief is given at the bottom of the environment. Moreover, the attitude type (belief, intention, or goal) is given on the far right bottom of the box. Though different attitude types are separated by environments, they can be nested so that agents can have beliefs, goals, and intentions about these attitudes. Figure 5 shows how meta-attitudes are used in VIEWGEN to represent the fact that the System believes that John intends to buy a car, but wants to convince him otherwise by getting him to believe that the car is a wreck (Lee and Wilks, 1996).

[INSERT FIGURE 5 HERE]

VIEWGEN avoids the use of shared or mutual beliefs. Rather, VIEWGEN attributes beliefs, goals, and intentions to other agents as required. This process is termed *ascription*. There are two methods of ascription: default ascription and stereotypical ascription. Default ascription applies to common attitudes which VIEWGEN assumes that any agent will hold and also ascribe to any other agent unless there is contrary evidence. This rule results in beliefs being pushed from outer belief environments to inner belief environments. For example, Figure 6 illustrates VIEWGEN ascribing beliefs to John about New Mexico. Stereotypical ascription is usually applied to “uncommon” attitudes which VIEWGEN assumes hold only for a particular class of agents.

[INSERT FIGURE 6 HERE]

In the VIEWGEN framework, locality and compatibility play an important role. Environments are thought of as means for modeling local beliefs, intentions, and goals. Every environment describes a set of attitudes from a certain point of view. Attitudes cannot be shared by different agents. All that an agent can do – before and during a dialogue with another agent – is to ascribe its own beliefs to the other agent unless there is contrary evidence. We can therefore consider VIEWGEN as an example of the compose-and-conquer approach. It is clear that the mechanism of ascription imposes particular relations between environments, and is possible only because the environments involved are supposed to be compatible. In the example depicted in Figure 6, after the belief ascription process, the environment containing John’s beliefs is compatible with the system’s beliefs only if, for every fact contained in the system’s beliefs, either this fact or its contrary is contained in the set of John’s beliefs.

3.2.2 *Representing belief*

The notion of context has been applied to formalize different aspects of intentional context, and, in particular, belief context. The approach we describe here was first introduced in a paper by Giunchiglia and Serafini (1994), and then used in several other papers to formalize different aspects of reasoning about belief. For example, Cimatti and Serafini (1995) use the notion of belief context introduced in this section to solve a well-known puzzle involving reasoning about belief and ignorance, namely the Three-Wise-Men problem. The work of Giunchiglia and his co-authors discusses the representation of an ideal and real reasoner using belief context (Giunchiglia et al., 1993; Giunchiglia and Giunchiglia, 1996). Benerecetti, Bouquet, and Ghidini (1998a) use belief context to solve the problem of the opaque and transparent reading of belief reports, while Fisher and Ghidini (1999) discuss the representation of resource-bounded deliberative agents. In the following, we call this approach

to the representation of belief *Hierarchical Belief* (HB).

HB focuses on a scenario with an agent ϵ who is acting in a world, who has beliefs about this world and is also observing and representing beliefs about a set $Ag = \{1, \dots, n\}$ of agents (possibly including itself). ϵ is also able to reason about its beliefs. Any element i in Ag is called an *agent index*. In order to formalize this scenario, HB introduces the notion of *belief context* (also called *view*) (Benerecetti et al. 1998b:65):

“A view [belief context] is a representation of a collection of beliefs that a reasoner [...] ascribes to an agent (including itself) under a given perspective.”

Examples of perspectives are: the beliefs that ϵ ascribes to itself, the beliefs that ϵ ascribes to another agent i , the beliefs that ϵ ascribes to an agent i about another agent j , and so on. As a convention, HB uses the Greek letter ϵ to indicate the belief context containing the beliefs that the reasoner ascribes to itself, and sequences of agent indexes (that is, sequences of elements of Ag) to label any other belief context. For instance, ϵi is the belief context containing the beliefs that ϵ ascribes to agent i from its perspective, and ϵij is the belief context containing the beliefs that ϵ ascribes to j from agent i 's perspective. Iterating the nesting, the belief context ϵijk formalizes the beliefs of agent i about j 's beliefs about k 's beliefs. Since, in the scenario presented above no confusion can arise, the prefix ϵ is often omitted. Figure 7 summarizes the idea underlying the HB approach, in a scenario in which the reasoner ϵ ascribes a collection of beliefs to three agents 1, 2, and 3.

[INSERT FIGURE 7 HERE]

HB formalization is based on the idea that a distinct language is associated with each belief context, and the interpretation of such a language is local to the belief context it is associated with. The idea is that a formula A in the external observer context, also written $\epsilon : A$ to stress the context dependence, expresses the fact that ϵ believes A . The same formula in context ijk , i.e. $ijk : A$, expresses the (more complex) fact that i believes that j believes that k believes that A (from the point of view of ϵ). HB are formalized using Local Models Semantics and MultiContext Systems. Each belief context is represented as a formal system $c_i = \langle L_i, \Omega_i, \Delta_i \rangle$ (see Section 2.2). To express statements about the world, L_i contains a set P_i of propositional constants. To express belief, L_i contains well formed formulae of the form $B_k("A")$, meaning that agent k believes the proposition expressed by A .

The other fundamental idea of the HB approach is that although distinct, the contents of different belief contexts are related. Relations between contexts, which in principle can be very different, express how the beliefs of an agent,

say agent ϵ , and the beliefs that ϵ ascribes to (itself or to) another agent, say agent i , are connected. A taxonomy of possible relations involving belief about belief is introduced by Giunchiglia and his co-authors in several papers (Giunchiglia et al., 1993; Giunchiglia and Giunchiglia, 1996). In this work the authors show that, depending on the relations between different contexts, the agent ϵ has different reasoning capabilities. An example of an obvious, and well studied, relation between belief contexts is the following: if a sentence of the form A is in ϵi , then a sentence of the form “ i believes that A ” is in ϵ . In this case, we say that ϵ is a *correct* observer (w.r.t. the sentence “ i believes that A ”). Another situation is when a sentence of the form A is in ϵi , only if a sentence of the form “ i believes that A ” is in ϵ . In this case, we say that ϵ is a *complete* observer (w.r.t. the sentence “ i believes that A ”). Formally, these relations between belief contexts are represented by the following *bridge rules* (Giunchiglia and Serafini, 1994):

$$\frac{\epsilon : B_i("A")}{\epsilon i : A} \mathcal{R}_{dn} \qquad \frac{\epsilon i : A}{\epsilon : B_i("A")} \mathcal{R}_{up}$$

As we can easily see from the discussion above, HB is a clear example of a compose-and-conquer use of context. Belief contexts are local theories, each of them expressing a partial and approximate representation of the world under a given perspective. Different relations, as in the one shown above, may exist among belief contexts. These relations allow the reasoner to (partially) compose the different belief contexts into a more comprehensive structure.

3.2.3 Recognizing mental states from communication

Representation of belief is naturally described by a compose-and-conquer approach, and the modeling of communication acts is a good example of this paradigm. The consequences of the speech acts are completely local and limited to the beliefs of an agent, but composition between different contexts is required in order to model changes that happen in the mental representations during communication acts. Moreover, it is not guaranteed that a unique, coherent theory exists, due to phenomena like deceit or misunderstanding.

The formal machinery of MultiContext Systems was exploited by Dragoni, Giorgini, and Serafini (2000) in order to give an account of the process of belief revisions in agents’ communication. In particular, it was used to model the consequences of an utterance in the mental states of the hearer. The authors adopt a plan-based vision of speech acts. They deal with the speech acts INFORM and REQUEST contained in most communication languages for artificial agents; INFORM and REQUEST are described in terms of pre-conditions and main effects in the set of beliefs. In particular, the authors focus their attention on agents with beliefs and intentions. They also assume the existence of a causal relationship between an agent’s mental state and the fact that the agent is possibly uttering a sentence. For example, one of the

causal relationships the authors assume for the speech act $\text{INFORM}(s,h,\phi)$, where s and h are agents and ϕ is a formula, is: (I1) s has the intention of bringing h into a mental state where ψ (in the general case $\psi \neq \phi$) is either believed or intended by h . As a consequence, abduction is used for updating the hearer's mental state from X to X' where X and X' are the sets of formulas that can be modified by receipt of the speech act. The next phase is devoted to intention recognition (Dragoni et al. 2000:131).

“By intention recognition we mean the hearer's ability to recognize the intention that induced the speaker to perform the speech act. Condition I1 states that a motivation for s to perform an $\text{INFORM}(s,h,\phi)$ is its intention of changing h mental states so that h believes or intends some new formula. To discover this intention, h checks the differences between its mental state before and after s executes $\text{INFORM}(s,h,\phi)$ (X and X' respectively), and then it revises X' to include the fact that s has the intention of causing this differences.”

Finally, the hearer updates its image of the speaker's mental state so that the speaker believes that his or her intentions have been satisfied. The work is related to work in the area of cognitive pragmatics (Airenti et al., 1993).

3.2.4 Agent's theory and applications

Contexts have been used for the specification of architectures for negotiating agents by Parsons, Sierra, and Jennings (1998). In a nutshell, the idea is to use the notion of context to represent the different components of an agent architecture, and to specify the interaction between the different components as appropriate rules between contexts. Here, the focus is on the use of contexts for designing modular architectures that are easy to maintain and to modify (1998:299).

“We use different contexts to represent different components of an agent architecture, and specify the interactions between the components by means of [...] rules between context. [...] This approach enforces a modular structure with well-defined interfaces, and thus accords well with good software practice.”

“This approach to system design affords all the traditional advantages of modularization in software engineering and enables complex artifacts to be designed out of simpler components.”

More in detail, an agent architecture consists of four components:

- *Units*: structural entities representing the main components of the architecture.
- *Logics*: declarative languages, each with a set of axioms and a number of rules of inference. Each unit has a single logic associated with it.
- *Theories*: sets of formulae written in the logic associated with a unit.
- *Bridge rules*: rules of inference which relate formulae in different units.

Units define the set of modules (contexts) of an agent architecture. *Logics* assign to each module a logic used to formally describe the content of the unit. *Theories* assign to each module a set of facts true in that module. Finally, *bridge rules* contain the set of rules that specify the interactions between modules. Notationally, bridge rules are written as follows:

$$\frac{u_1 : \phi \quad u_2 : \psi}{u_3 : \theta}$$

This particular bridge rule means that the formula θ may be inferred in unit u_3 because of the fact that ϕ and ψ are derivable in the units tagged with u_1 and u_2 , respectively. Figure 8 shows an example of architecture, where the units are u_1 , u_2 , u_3 , and c . In this architecture, u_1 is formalized as a propositional logic, while u_2 , u_3 , and c are formally described using a first order logic. Bridge rules are depicted as arcs connecting the units. As a more practical example, units u_1 , u_2 , u_3 can contain the beliefs, desires, and intentions of the agent, respectively, while c is the communication unit which is responsible for enacting the agent’s communication needs. By imposing the appropriate bridge rules between the beliefs, desires, and intentions units, Parsons and his co-authors show how to model different agent’s behaviors. The formal framework underlying the agent architecture defined by Parsons, Sierra, and Jennings is that of MultiContext Systems.

[INSERT FIGURE 8 HERE]

It is disputable whether the notion of context presented in this work follows a divide-and conquer or a compose-and-conquer approach. The authors do not state explicitly either the fact that a context is an expression of a viewpoint or the existence of a unique model as a primitive concept.

3.3 Knowledge and data integration

Knowledge and data integration is another area where both the divide-and-conquer and the compose-and-conquer approaches to contexts have been widely applied. The difference between the two different approaches is crucial. The divide-and conquer use of context in knowledge and data integration is based on the fact that a unique global schema can always be reconstructed. As a consequence, the semantic heterogeneity between different information sources (represented as different contexts) can always be resolved. The compose-and-conquer use of context, on the contrary, aims at providing formal systems for a federation of heterogeneous data or knowledge bases, possibly developed independently. Each knowledge base can be seen as a set of views of an ideal database which is often impossible or very complex to reconstruct completely.

In this respect, the main goal of the compose-and-conquer approach is to (partially) relate semantically heterogeneous information sources and not to integrate such information into a unique and homogeneous schema.

In the following, we present several uses of context in data integration. The main problems we consider are the integration of different information sources and the partitioning of very large knowledge bases.

3.3.1 *Integration of different information sources*

Context based frameworks have been used to provide formal models for the integration of information (or knowledge) coming from different sources, often developed independently. Their use is based on the intuition that different information sources integrated in a unique system (or federation) can be thought of as partial views (thought of as contexts) on a common world. Often the different information sources have very little in common. They are obviously distributed, that is, each information source is part of a different system and contains a specific piece of knowledge. They are redundant, meaning that the same piece of knowledge may be represented, possibly from different perspectives, in more than one information source. They are partial, that is, the information contained in an information source may be incomplete. Finally, they are autonomous, that is, each information source has a certain degree of autonomy regarding the design, the execution, and the communication with the other databases. As a consequence, information sources may adopt different conceptual schemata (including domains, relations, naming conventions, ...), and certain operations are performed locally by the information source, without interactions with the others. It is therefore easy to consider the knowledge (data) contained in each information source as context dependent. Apart from this initial common assumption, the compose-and-conquer and divide-and-conquer approaches address different problems and provide different solutions.

3.3.1.1 The compose-and-conquer approach. (Ghidini and Serafini, 1998a,b) use Local Models Semantics as a formal framework for information integration. In (1998b:192), the following example is discussed:

“Let m be a mediator of an electronic market place for fruits, composed of three fruits sellers: 1, 2, and 3. m collects information about fruit prices from 1, 2, and 3 and integrates it in a unique homogeneous database. Customers that need information about fruit prices, instead of connecting each seller, can submit a single query to the mediator.”

They describe a formalization of the exchange of information in the example above by means of four contexts (a context for each fruit seller and one for the mediator) and the appropriate connections between the four contexts.

The representation of the different contexts in this example is done by using Local Models Semantics. It associates a different (set of) first order model(s)

to each database. This enables the authors to formalize the fact that each database is associated with a specific domain. For instance the sellers can sell different subsets of fruits, and therefore the domains of their databases can differ. Moreover, it enables the framework to represent the fact that the domain of fruits can be represented at different levels of detail by different sellers. For instance, database 1 may contain prices for red apples and yellow apples, while database 2 and 3 abstract the dependence on the color away and do not make this distinction. Finally, Ghidini and Serafini point out that associating different models to different databases enables the framework to capture the fact that prices for different sellers might be not homogeneous, depending on their particular viewpoint. For instance, they assume that prices of fruits in database 1 don't include taxes, while they do in databases 2, 3 and m .

Assigning different models to the different databases enables the authors to formalize the differences between each database. The next step in order to meaningfully integrate knowledge coming from the different databases is to carefully consider extra information that is left implicit in the representation of knowledge itself. In this example, this consists in

- (i) the differences between the domains of the representation; and
- (ii) the different interpretations of the predicate $costs(x, y)$ (meaning that fruit x costs y Euro) in the different databases, depending on the fact that prices do or do not include taxes.

According to Ghidini and Serafini (1998a), the relations between the different domains of the representation are represented by introducing *domain relations*, i.e., relations between the interpretation domains of the different databases. A domain relation may, for instance, relate a “more abstract” object (e.g. apple) in the domain of a database to a set of “less abstract” objects (e.g. red–apple, green–apple) in the domain of another database. In the same paper, different perspectives on related information (e.g. the different perspective on $costs(x, y)$ in our example) are represented by using *interpretation constraints*. An interpretation constraint is a relation between formulae contained in the languages of the different databases. For instance the different (but related) meaning of the predicate $costs(x, y)$ in database 1 and in the database m are represented by using the following expression:

$$1 : costs(x, y) \rightarrow m : \exists y' costs(x, y') \wedge y' = 1.07 * y$$

Its meaning is that every time the models of database 1 satisfy the formula

$$costs(x, y)$$

then the models of the mediator database must satisfy the formula

$$\exists y' costs(x, y') \wedge y' = 1.07 * y$$

which means that item x has price y' which is obtained by adding the correct amount of taxes to y .

As we can see from the example, the goal here is not to integrate the four databases in a unique schema, described with the same language, and interpreted over the same domain. On the contrary, Local Models Semantics aims at meaningfully relating knowledge coming from the different databases.

3.3.1.2 The divide-and-conquer approach. Farquhar, Dappert, Fikes, and Pratt (1995), on the other hand, use Buvač and Mason’s logic of context to provide a framework for information source integration. The main idea is to represent the information contained in each information source (database) by means of two contexts (1995):

“The information source context is a direct translation of a database schema into logic *without* resolving semantic conflicts [...]. The semantic context holds the translation with the semantic conflicts resolved.”

In order to integrate the different information sources, an integrating context containing axioms that lift from several semantic contexts is added, which provides a global schema of the integrated system.

The work by Farquhar and his co-authors is a prototypical example of the divide-and-conquer approach. Semantic contexts are used to solve semantic conflicts between databases so that the global model (schema) of the world can be reconstructed in the integrating context.

3.3.2 Partitioning knowledge bases

A context-based approach to the problem of specifying redundancy among different databases while maintaining an high degree of autonomy has been proposed by Mylopoulos and Motschnig-Pitrik (1995). In this paper, the authors describe a set of criteria for splitting a database into a set of (possibly overlapping) partitions. This work is particularly relevant, as it provides mechanisms for the management of different overlapping partitions based on the notion of (possibly overlapping) *context*.

According to Mylopoulos and Motschnig-Pitrik (1995), an information base is composed of units. A unit might represent an entity, object, attribute, relationship, rule, method, etc. A context is a special unit representing the decomposition of an information base in which they appear (Mylopoulos and Motschnig-Pitrik 1995:46):

“The definition of each context includes three components which define respectively the *contents* of the context, the local names (*lexicon*) used for units in the context’s contents and authorization rules for combinations of different users and transactions:

$$P(\text{Units}) \times (\text{Identifiers} \leftarrow \text{Units}) \times \text{Predicates} \supseteq \text{Contexts}$$

That is, a context c is a triple where:

- the first element is a set of units (i.e., an element of $P(\mathbf{Units})$) which define the content of the context.
- the second element is a mapping from $\mathbf{Identifiers}$ to \mathbf{Units} (i.e., an element of $\mathbf{Identifiers} \leftarrow \mathbf{Units}$) which defines the lexicon of the context.
- the third element is a set of formulae which determine whether a certain user is authorized to execute a certain transaction in the context.

The notion of context represents the fact that a set of units can be partitioned into a set of different and partially overlapping modules, and that each of the modules can assign different Identifiers to the units and have different transaction rules. Each context has its own content, lexicon, and transaction rules but, although named with different identifiers, the same unit can belong to two (or more) contexts. In fact, Mylopoulos and Motschnig-Pitrik are interested in establishing mechanisms for change propagation: these are mechanisms that establish whether the effects of a change operation performed on a certain unit in a certain context are (or are not) visible in other contexts (1995:50).

“The effects of a change operation performed on a certain unit, say o , with respect to one context are not automatically visible in other contexts which also contain o . The visibility of a change in other contexts depends on whether that change is *propagated*.

In particular, change propagation between two contexts is declared through two complementary operations

[...] Once these operations have been executed, we will say that there exists a *propagation channel*.”

The work described by Mylopoulos and Motschnig-Pitrik has strong connections with the divide-and-conquer approach. The same problem of partitioning knowledge bases seems to impose the existence of a global model of the world (the initial knowledge base), and uses context as a device for focusing on (smaller, simpler) portions of this model. Nonetheless, the mechanism of propagation channels can be seen as a mechanism for diversifying the different partitions so that they can express some partial viewpoint on this general model of the world. In fact, although different contexts may in principle overlap, only the existence of a propagation channel makes this overlapping count. If no propagation channel is established, the overlapping is totally irrelevant, as the changes made to units which are in a context can't be propagated to other contexts containing the same unit.

3.3.3 Very Large Knowledge Bases: CYC

CYC is an attempt to build a massive knowledge base so that one (person or machine) can apply it to some reasoning mechanism. The importance of building very large knowledge bases as one of the steps in developing programs with some sort of intelligence is well motivated by Lenat et al. (1990). In the

following quotation, the author argues that the lack of very large knowledge bases covering a wide area of (human) knowledge is one of the reasons why “expert systems” failed in their attempt to be regarded as intelligent programs (Lenat et al. 1990:32):

“Suppose an expert system has the following four rules:

IF frog(x), THEN amphibian(x)

IF amphibian(x), THEN laysEggsInWater(x)

IF laysEggsInWater(x), THEN livesNearLotsOf(x, Water)

IF livesNearLotsOf(x, Water), THEN \neg livesInDesert(x)

Given the assertion frog(Freda), those rules could be used to conclude that various facts are true about Freda: amphibian(Freda), laysEggsInWater(Freda), \neg livesInDesert(Freda), etc. Yet the program would not “know” how to answer questions like: Does Freda lay eggs? Is Freda sometimes in water?

Humans can draw not only those direct conclusions from laysEggsInWater(Freda), but can also answer slightly more complex queries which require a modicum of “outside” knowledge: Does Freda live on the sun? Was Freda born live or from an egg? Is Freda a person? [...]

Carefully selecting just the fragment of relevant knowledge leads to adequate but “brittle” performance: when confronted by some unanticipated situation, the program is likely to reach the wrong conclusion.”

Unfortunately, the solution of building a knowledge base composed of a million rules has several drawbacks (Lenat 1999:3):

“During the 1984-1989 time period, as the CYC common sense knowledge base grew ever larger, it became increasingly difficult to shoehorn every fact and rule into the same flat world. Finally, in 1989, as CYC exceeded 100,000 rules in size, we found it necessary to introduce an explicit context mechanism. That is, we divided the KB up into a lattice of hundreds of contexts, placing each CYC assertion in whichever context(s) it belonged.”

Contexts in CYC have a fine internal structure. Lenat (Lenat, 1999) identifies a dozen mostly-independent dimensions along which contexts vary. Each region of this 12-dimensional space implicitly defines a context. The capability of importing an assertion from one context into another is provided by lifting assertions similar to the ones described in Section 3.1.

Differently from Mylopoulos and Motschnig-Pitrik (1995), CYC entirely follows the divide-and-conquer approach. Contexts are used precisely for partitioning the CYC knowledge base, which provides the global model of the world, into (smaller, simpler) portions.

4 Conclusions

In this paper, we have suggested that there are two types of theory of context in KRR: the first, which we called the *divide-and-conquer* theory, sees context as a way of partitioning a global model of the world into smaller and simpler pieces; the second, which we called the *compose-and-conquer* theory, sees context as a local theory of the world in a network of relations with other local theories. We have discussed the ins and outs of the theories of the two types, and have shown that each type of theory leads quite naturally to the use of context for addressing different issues in KRR, or for providing conceptually different solutions to the same issues.

As a final remark, we would like to note that there seems to be a crucial difference between the notion of context in KRR and the notion of context used in disciplines like philosophical logic, cognitive psychology, and linguistics. In KRR, context is thought of as an agent's partial representation of the world (that is, something similar to a theory), and therefore is assumed to be related to the cognitive state of the agent. In many other disciplines, context is conceived of as a collection of features of the location (in a broad sense) in which an agent produces a linguistic expressions (or a thought), and is therefore assumed to be related to the state of the world. This difference is very apparent even when we observe that the prototypical problem of context in KRR has to do with different possible representations of a given situation, whereas in the other disciplines we mentioned, it is the problem of indexicality, which is mostly ignored in KRR.

This tension between a cognitive and a metaphysical notion of context is an example of the dichotomies that appear in many theories of context. Other instances are: subjective vs. objective, internal vs. external, cognitive vs. social, and so on. In our view, the implication of this is that we are still in need of a general foundation for a theory of context that can account for the different uses made of context in different disciplines to address different problems. In (Benerecetti et al., 2000) there is a first attempt to define the boundaries of such a theory, but the way is still very long.

References

- Airenti, Gabriella, Bara, Bruno G., Colombetti, Marco, 1993. Conversation and behaviour games in the pragmatics of dialogue. *Cognitive Science* 17, 197–253.
- Akman, Varol, Bouquet, Paolo, Thomason, Rich, Young, Roger A. (Eds.), 2001. Proceedings of the Third International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT'2001). Vol. 2116 of Lecture Notes in Artificial Intelligence. Springer Verlag.
- Akman Varol, Surav, Mehmet, 1996. Steps toward formalizing context. *AI Magazine* 17 (3), 55–72.
- Attardi, Giuseppe, Simi, Maria, 1995. A formalisation of viewpoints. *Fundamenta Informaticae* 23 (2–4), 149–174.
- Ballim, Afzal, Wilks, Yorick, 1991. Artificial believers: The ascription of belief. Lawrence Erlbaum, Hillsdale, NJ.
- Benerecetti, Massimo, Bouquet, Paolo, Ghidini, Chiara, 1998a. Formalizing Belief Reports – The Approach and a Case Study. In: Giunchiglia, F. (Ed.), Proceedings of the 8th International Conference on Artificial Intelligence, Methodology, Systems, and Applications (AIMSA'98). Vol. 1480 of Lecture Notes in Artificial Intelligence. Springer Verlag, pp. 62–75.
- Benerecetti, Massimo, Bouquet, Paolo, Ghidini, Chiara, 2000. Contextual Reasoning Distilled. *Journal of Theoretical and Experimental Artificial Intelligence* 12 (3), 279–305.
- Benerecetti, Massimo, Giunchiglia, Fausto, Serafini, Luciano, 1998b. Model Checking Multiagent Systems. *Journal of Logic and Computation, Special Issue on Computational & Logical Aspects of Multi-Agent Systems* 8 (3), 401–423.
- Bouquet, Paolo, Giunchiglia, Fausto, 1995. Reasoning about Theory Adequacy. A New Solution to the Qualification Problem. *Fundamenta Informaticae* 23 (2–4), 247–262.
- Bouquet, Paolo, Serafini, Luciano, 2001. Two formalizations of context: a comparison. In: Akman et al. (2001), pp. 87–101.
- Bouquet, Paolo, Serafini, Luciano, Brézillon, Patrick, Benerecetti, Massimo, Castellani, Francesca (Eds.), 1999. Proceedings of the Second International and Interdisciplinary Conference on Modelling and Using Context (CONTEXT'99). Vol. 1688 of Lecture Notes in Artificial Intelligence. Springer Verlag.
- Brézillon, Patrick, 1999. Context in problem solving: A survey. *The Knowledge Engineering Review* 14 (1), 1–34.
- Buvač, Saša, 1996. Quantificational logic of context. In: Proceedings of the Thirteenth American National Conference on Artificial Intelligence (AAAI-96).
- Buvač, Saša, Mason, Ian A., 1993. Propositional logic of context. Proc. of the 11th American National Conference on Artificial Intelligence (AAAI-93), 412–419.

- Cimatti, Alessandro, Serafini, Luciano, 1995. Multi-Agent Reasoning with Belief Contexts: the Approach and a Case Study. In: Wooldridge, M., Jennings, N. R. (Eds.), *Intelligent Agents: Proceedings of 1994 Workshop on Agent Theories, Architectures, and Languages*. No. 890 in *Lecture Notes in Computer Science*. Springer Verlag, pp. 71–85.
- Dinsmore, John, 1991. *Partitioned Representations*. Kluwer Academic Publishers.
- Dragoni, Aldo F., Giorgini, Paolo, Serafini, Luciano, 2000. *Mental State Recognition from Communication*. Tech. Rep. 0007-03, ITC-IRST, Trento, Italy.
- Farquhar, Adam, Dappert, Angela, Fikes, Richard, Pratt, Wanda, 1995. Integrating Information Sources Using Context Logic. In: *Proceedings of AAAI Spring Symposium on Information Gathering from Distributed Heterogeneous Environments*.
- Fauconnier, Gilles, 1985. *Mental Spaces: aspects of meaning construction in natural language*. MIT Press.
- Fisher, Michael, Ghidini, Chiara, 1999. Programming Resource-Bounded Deliberative Agents. In: *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI'99)*. Morgan Kaufmann Publ., Inc, pp. 200–206.
- Gabbay, Dov M., Nossum, Rolf T., 1997. Structured Contexts with Fibred Semantics. In: *Proceedings of the 1st International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT-97)*. Rio de Janeiro, Brazil, pp. 46–55.
- Ghidini, Chiara, 1999. Modelling (Un)Bounded Beliefs. In: Bouquet et al. (1999), pp. 145–158.
- Ghidini, Chiara, Giunchiglia, Fausto, 2001. Local Models Semantics, or Contextual Reasoning = Locality + Compatibility. *Artificial Intelligence* 127 (4), 221–259.
- Ghidini, Chiara, Serafini, Luciano, 1998a. Distributed First Order Logics. In: Gabbay, D., de Rijke, M. (Eds.), *Frontiers Of Combining Systems 2. Studies in Logic and Computation*. Research Studies Press.
- Ghidini, Chiara, Serafini, Luciano, 1998b. Information Integration for Electronic Commerce. In: Sierra, C., Noriega, P. (Eds.), *Proceedings of the Workshop on Agent Mediated Electronic Trading (AMET'98)*. Vol. 1571 of *Lecture Notes in Artificial Intelligence*. Springer Verlag, pp. 189–206.
- Ghidini, Chiara, Serafini, Luciano, 1998c. Model Theoretic Semantics for Information Integration. In: Giunchiglia, F. (Ed.), *Proceedings of the 8th International Conference on Artificial Intelligence, Methodology, Systems, and Applications (AIMSA'98)*. Vol. 1480 of *Lecture Notes in Artificial Intelligence*. Springer Verlag, pp. 267–280.
- Giunchiglia, Enrico, Giunchiglia, Fausto, 1996. Ideal and Real Belief about Belief. In: *Practical Reasoning, International Conference on Formal and Applied Practical Reasoning (FAPR'96)*. No. 1085 in *Lecture Notes in Artificial Intelligence*. Springer Verlag, pp. 261–275.
- Giunchiglia, Fausto, 1993. Contextual reasoning. *Epistemologia*, XVI, 345–

- Giunchiglia, Fausto, Giunchiglia, Enrico, Costello, Tom, Bouquet, Paolo, 1996. Dealing with Expected and Unexpected Obstacles. *Journal of Experimental and Theoretical Artificial Intelligence* 8, 173–190.
- Giunchiglia, Fausto, Serafini, Luciano, 1994. Multilanguage hierarchical logics (or: how we can do without modal logics). *Artificial Intelligence* 65, 29–70.
- Giunchiglia, Fausto, Serafini, Luciano, Giunchiglia, Enrico, Frixione, Marcello, 1993. Non-Omniscient Belief as Context-Based Reasoning. In: Proc. of the 13th International Joint Conference on Artificial Intelligence. Chambéry, France, pp. 548–554.
- Guha, Ramanathan V., 1991. Contexts: a Formalization and some Applications. Tech. Rep. ACT-CYC-423-91, MCC, Austin, Texas.
- Kokinov, Boicho, 1995. A Dynamic Approach to Context Modelling. In: Brezillon, P., Abu-Hakima, S. (Eds.), Working Notes of the IJCAI-95 Workshop on “*Modelling Context in Knowledge Representation and Reasoning*”. Montreal (Canada).
- Lee, Mark, Wilks, Yorick, 1996. An Ascription-based Approach to Speech Acts. In: Proceedings of the 16th International Conference on Computational Linguistics (COLING-96).
- Lenat, Douglas B., Sierra, Carles, 1999. The Dimensions of Context Space. Tech. rep. CYCorp, <http://www.cyc.com/context-space.rtf,doc,txt>.
- Lenat, Douglas B., Sierra, Carles, Guha, Ramanathan V., Pittman, Karen, Pratt, Dexter, Shepherd, Mary, August 1990. Cyc: toward programs with common sense. *Communications of the ACM* 33 (8), 30–49
- McCarthy, John, 1987. Generality in Artificial Intelligence. *Communications of ACM* 30 (12), 1030–1035.
- McCarthy, John, 1993. Notes on Formalizing Context. In: XIII International Joint Conference on Artificial Intelligence. Chambéry (France), pp. 555–560.
- Mylopoulos, John, Motschnig-Pitrik, Renate, 1995. Partitioning Information Bases with Contexts. In: Third International Conference on Cooperative Information Systems. Vienna.
- Noriega, Pablo, Sierra, Carles, 1996. Towards Layered Dialogical Agents. In: Müller, J. P., Wooldridge, M. J., Jennings, N. R. (Eds.), Third International Workshop on Agent Theories, Architectures, and Languages (ATAL-96). Vol. 1193 of Lecture Notes in Artificial Intelligence. Springer Verlag.
- Parsons, Simon, Sierra, Carles, Jennings, Nick R., 1998. Agents that reason and negotiate by arguing. *Journal of Logic and Computation* 8 (3), 261–292.
- Sperber, Dan, Wilson, Deirdre, 1986. *Relevance: Communication and Cognition*. Basil Blackwell.
- Thomason, Richmond, 1999. Type theoretic foundation for context, part 1: Contexts as complex type-theoretic objects. In: Bouquet et al. (1999), pp. 351–360.
- Weyhrauch, Richard W., 1980. Prolegomena to a Theory of Mechanized Formal Reasoning. *Artificial Intelligence* 13 (1), 133–176.

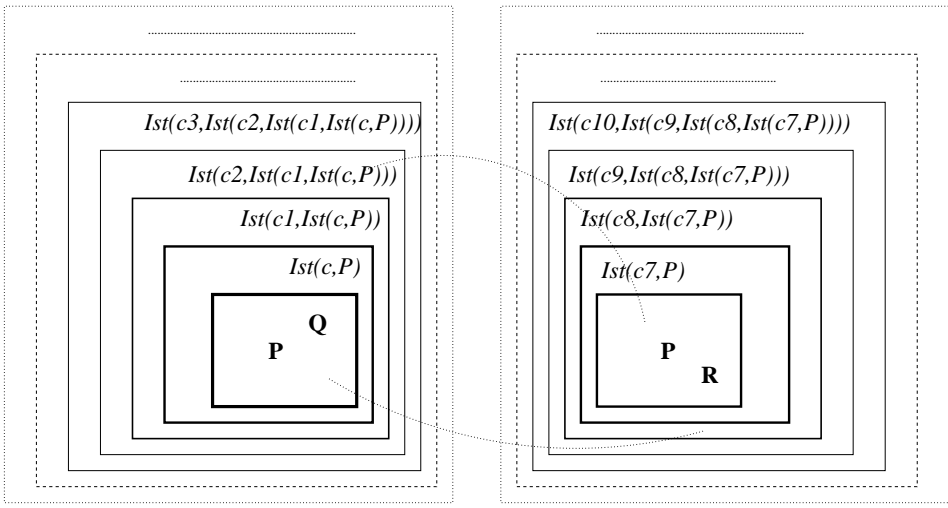


Fig. 1. A divide-and-conquer theory of context: LoC

- (PL) $\vdash_{\bar{c}} \phi$ If ϕ is an instance of a classical tautology
- (K) $\vdash_{\bar{c}} \text{ist}(c, p \supset q) \supset \text{ist}(c, p) \supset \text{ist}(c, q)$
- (Δ) $\vdash_{\bar{c}} \text{ist}(c_1, \text{ist}(c_2, p) \vee q) \supset \text{ist}(c_1, \text{ist}(c_2, p)) \vee \text{ist}(c_1, p)$
- (MP) $\frac{\vdash_{\bar{c}} p \quad \vdash_{\bar{c}} p \supset q}{\vdash_{\bar{c}} q}$
- (CS) $\frac{\vdash_{\bar{c}} p}{\vdash_{\bar{c}} \text{ist}(c, p)}$

Fig. 2. Axioms and inference rules for PLC

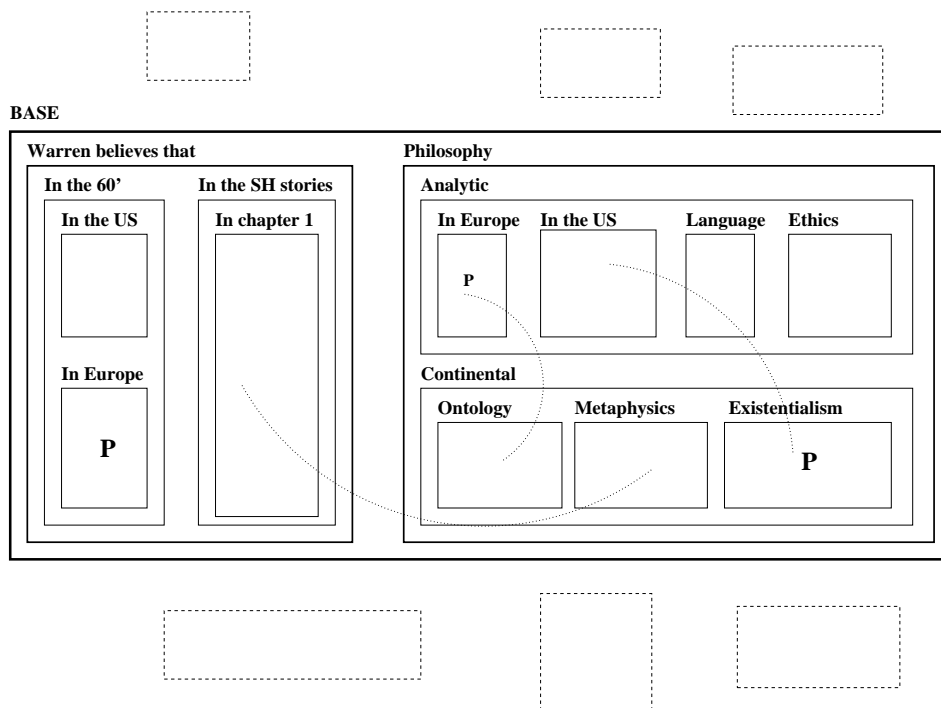


Fig. 3. Dinsmore's partitioned representations

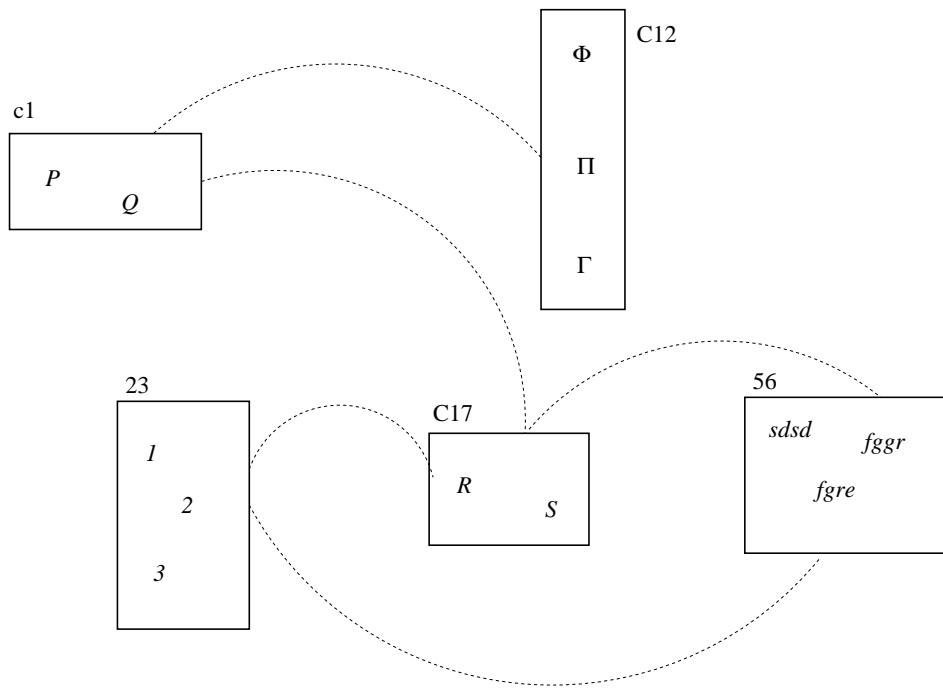


Fig. 4. A compose-and-conquer theory of context: LMS.

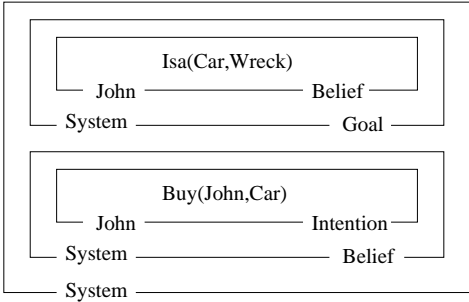


Fig. 5. Meta-attitudes in VIEWGEN.

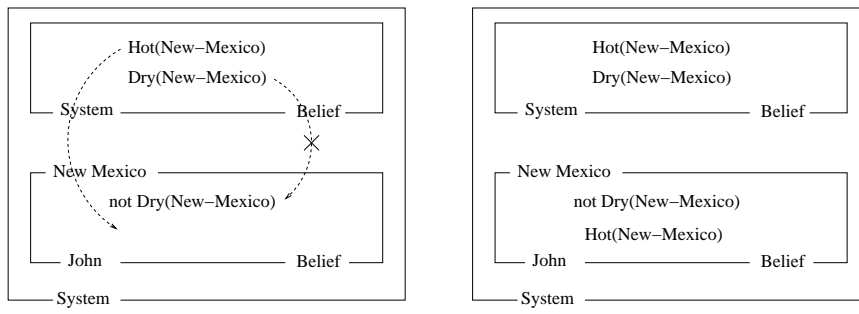


Fig. 6. VIEWGEN: before ascription(left) and after ascription (right).

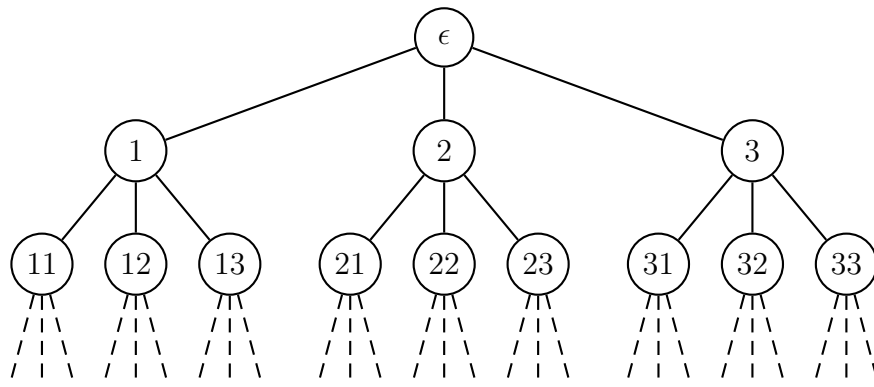


Fig. 7. The context structure of beliefs in a scenario with three agents.

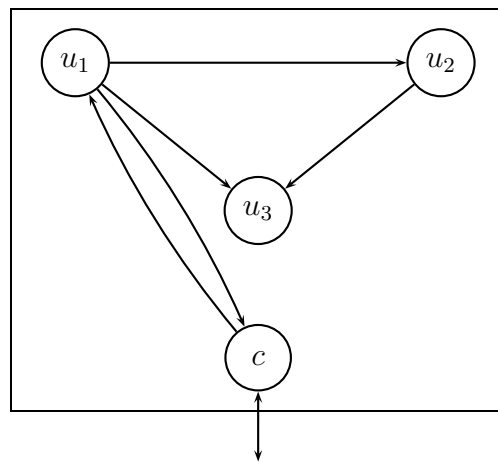


Fig. 8. A context-based agent architecture.