Non-omniscient belief as context-based reasoning*

Fausto Giunchiglia[1,2] Luciano Serafini[1] Enrico Giunchiglia[2] Marcello Frixione[2]
Mechanized Reasoning Group
[1]RST, 38050 Povo, Trento, Italy
[2]DIST - University of Genoa, Via Opera Pia 11 A, Genova, Italy
phone:+39 461 814440
{fausto,serafini}@irst.it {enrico,frix}@dist.unige.it

Abstract

This paper describes a general framework for the formalization of monotonic reasoning about belief in a multiagent environment. The agents* beliefs are modeled as logical theories. The reasoning about their beliefs is formalized in still another theory, which we call the theory of the computer. The framework is used to model non-omniscient belief and shown to have many advantages. For instance, it allows for an exhaustive classification of the "basic" forms of non logical omniscience and for their "composition" into the structure of the system modeling multiagent omniscient belief.

1 The approach

This paper describes a general framework for the formalization of monotonic reasoning about belief in a multiagent environment. The most common solution is to take a first order (propositional) theory, to extend it using a set of modal operators, $\{B^i\}_{i\in I}$, and to take B^iA as meaning that an agent a_i believes A (see for instance [Halpern and Moses, 1985]). There is only one theory of the world, however this theory proves facts about the agents' beliefs. According to a first interpretation, this theory is taken to model things how they really are. It is therefore a finite (and possibly incomplete) presentation of what is true in the world, and the fact that $B^i A$ is a theorem means that it is, in fact, the case that a^i believes A. According to another interpretation, this theory is taken to be the perspective that a generic reasoner has of the world. It is therefore a finite presentation of the reasoner's beliefs, and the fact that B^iA is a theorem means that the reasoner believes that a_i believes A.

Once one accepts the second interpretation (as we do), a mechanized theory is naturally taken as representing the beliefs of the computer where it is implemented. Moreover, in the case of multiagent belief, a further step is to have, together with the theory of the computer, one theory (at least, see later) for each agent.

*Alessandro Cimatti and Kurt Konolige have provided very useful feedback and suggestions. The work at IRST has been done as part of the MAIA project. This paper is a short version of the IRST technical report #9206-03.

The theory of the computer plays the same role as the unique theory in the modal logics approach. The agents' theories are the (mental) representations that the computer has of the agents themselves. The computer has beliefs about the beliefs of the agents not because, as it happens in the single theory approaches, it "simulates" them in its own theory but, rather, because it can infer, say, $B^i("A")$ from the fact that A is a theorem in the a_i's theory. (In this framework, theories are called "contexts" and the formal systems, which are basically defined as sets of contexts, are called *multicontext systems* (MC systems) and, sometimes, *multtlanguage systems* (ML systems) [Giunchiglia, 1991; Giunchiglia, 1993].)

All the previous work on the formalization of propositional attitudes, namely modal logics and the syntactic approach (where belief is a first order predicate, see for instance [Konolige, 1982]) makes use of a single theory. However, ideas similar to ours, which suggest the use of multiple distinct theories, have been exploited in much applied work in computational linguistics and AI (see for instance [Wilks and Biem, 1979; Giunchiglia and Weyhrauch, 1988]). There are in fact many irnplementational advantages, all deriving from the fact that reasoning becomes intrinsically localized [Giunchiglia, 1993] (for instance: the modularization of the knowledge base, computational efficiency ~ due to the division of the search space in smaller search spaces - and the possibility of parallelization of the reasoning process). Our formalism seems quite close, arguably closer than the previous approaches, to the current practice in the more applied irnplementational work. Indeed this is one of the main motivations underlying our work and MC systems have been implemented inside the GETFOL system, [Giunchiglia, 1992] an extension of a reimplementation of the FOL system [Weyhrauch, 1980]. However in this paper we focus on the representational issues and argue that our approach allows for more natural and more intuitive formalizations (see also [Dinsmore, 1991]).

In [Giunchiglia and Serafini, 1991] we have described and motivated the basic system, which models logically omniscient agents. In this paper we take a step further and treat the problem of non saturated belief, that is the fact that real agents never believe all the logical consequences of their basic beliefs. Our treatment is based on a distinction between ideal and real reasoners. We show how ideal and (different) real reasoners can be modeled by contexts with different structural properties (section 2) and/or different connections among them (section 3). Our approach presents various advantages. First, it allows for an exhaustive classification of all the basic forms of non logical omniscience (see sections 2, 3). Second, all the various forms of non logical omniscience can be identified as a distinct structural property of the system. Third, complex forms of non logical omniscience can be "constructed" by composing simpler forms into the structure provided by the system for saturated belief (see sections 2, 3). All the cases studied in the past plus new interesting ones are captured by this classification (see section 4). Fourth, deductions can be performed very naturally by exploiting the fact that each reasoner is modeled as a distinct theory (section 5)[1].

2 Ideal and real reasoners

As informally described in the previous section, reasoners (*e.g.* agents and the computer itself) are modeled as logical theories, which we present as axiomatic formal systems. Technically, we define a formal system Σ as a pair consisting of a language L_Σ and a set of theorems $T_\Sigma \subseteq L_\Sigma$, *i.e.* $\Sigma = \langle L_\Sigma, T_\Sigma \rangle$. L_Σ is defined as the smallest set generated from the set of propositional letters (atomic formulas) P_Σ and closed (by C) under the set of formula building operations W_Σ, in formulas $L_\Sigma = C(P_\Sigma, W_\Sigma)$. Analogously, T_Σ is defined as the smallest set generated from the set of axioms Ω_Σ and closed under the set of inference rules Δ_Σ, in formulas $T_\Sigma = C(\Omega_\Sigma, \Delta_\Sigma)$. Ω_Σ is also called the set of the basic beliefs.

A first approach is to take a reasoner $I = \langle L_I, T_I \rangle$, such that, given a set P_I of atomic formulas and a set $\Omega_I \subseteq L_I$ of true facts, it believes all the tautologies and logical consequences of Ω_I. Ω_I is therefore I's set of basic beliefs. Technically, this implies that W_I and Δ_I are complete for a propositional language and for propositional logic, respectively. In this case we say that I is an *ideal, logically omniscient reasoner*.

The problem is that real reasoners are not logically omniscient. Given them enough knowledge and resources (*e.g.* space, time), real reasoners tend to converge to the behaviour of an idealized reasoner but this may never be the case. Thus, if we take $E = \langle L_E, T_E \rangle$ to be a reasoner which is the "realized" version of I, a plausible model for E is that $L_E \subseteq L_I$ and $T_E \subseteq T_I$. This captures the intuition that the ideal reasoner is the limit *saturated* case of the real reasoner in the sense that E's beliefs are always a subset of I's beliefs. However it is possible to ask for stronger notions of realization. For instance, if $\vdash_\Sigma$ is the consequence relation of Σ (that is the set of pairs (Γ, A), with $\Gamma \cup \{A\} \in L_E$, such that there exists a derivation of A from Γ), we could impose that $\vdash_E \subseteq \vdash_I$. Intuitively this means that not only do we ask that E's beliefs are a subset of I's but also that E's ability to compute the logical consequences of sets of assumptions is weaker than I's. In the following we consider the case $T_E \subseteq T_I$.

The recursive definition of L_Σ and T_Σ allows for an exhaustive classification of all the possible ways for a reasoner to be not ideal. We may have $L_E \subset L_I$ and $T_E \subset T_I$ because E's basic sets (P_E, Ω_E) or because E's constructors (W_E for formulas, Δ_E for derivations and theorems) are "incomplete". Technically, this intuitive notion of incompleteness can be made precise as follows.

Definition 2.1 *Let $S_1 = C(B_1, F_1)$, $S_2 = C(B_2, F_2)$ be two sets such that $S_1 \subset S_2$. We say that*

- B_1 *is* incomplete *wrt B_2, and write $B_1 \prec B_2$, iff $C(B_1, F_1) \subset C(B_2, F_1)$. We also say that S_1 is incomplete (wrt S_2) in the basic facts.*
- F_1 *is* incomplete *wrt F_2, and write $F_1 \prec F_2$, iff $C(B_1, F_1) \subset C(B_1, F_2)$. We also say that S_1 is incomplete (wrt S_2) in the construction rules.*

"$\preceq$" is defined in the obvious way. Set equality, written "=", is the equivalence relation for $\preceq$. L_E and T_E may be incomplete (wrt L_I and T_I respectively) because of their basic facts or because of their construction rules (or both). This suggests the classification below. Let us start from the incompleteness in the language.

Incompleteness in the signature ($P_E \prec P_I$). In this case we take $P_E \prec P_I$ to mean $P_E \subset P_I$, namely that E is not aware of some primitive propositions (we suppose that real reasoners know the logical connectives). As pointed out in [Fagin and Halpern, 1988], supposing that P is the proposition expressing that personal computer prices are going down, it is very likely that a Bantu tribesman is not aware of P.

Incompleteness in the formation rules ($W_E \prec W_I$). E is aware of a set of propositions. Still he does not succeed in composing some complex propositions. This may happen for two reasons. In the first case, similarly to the case of incompleteness in the signature, E simply does not know some formation rules ($W_E \subset W_I$). For instance even if it knows the logical connective for conjunction, he does not know how to build $P_1 \wedge P_2$ from P_1 and P_2. In the second case E suffers from a form of limitation of resources and it is not able to form sentences beyond a certain level of completeness, *e.g.* longer than a certain limit. For instance a sentence with one hundred constants rises problems of space and time. Notice that in this case ($W_E \not\subset W_I$).

A similar classification can be given for T_E.

Incompleteness in the axioms ($\Omega_E \prec \Omega_I$). This means $C(\Omega_E, \Delta_E) \subset C(\Omega_I, \Delta_E)$. A first possibility is that $\Omega_E \subset \Omega_I$. This may have two causes. The first is due to the fact that incompleteness in the signature or in the formation rules may imply $\Omega_E \subset \Omega_I$ (as $\Omega_E \subseteq L_E$). However, and this is the second case, we may think of situations where E is aware of a proposition but without knowing whether it holds. Technically these two situations are distinguished as in the second case we have

[1] For lack of space, these ideas are described only partially. For instance, we state theorems but we do not prove them. We do not consider nested belief. We do not give semantics ([Giunchiglia *et* al., 1992] presents a somewhat old semantics for the basic saturated case). We do not give any complexity argument; although it is intuitive that reasoning with contexts saves time, we have not done any in depth analysis on this issue.

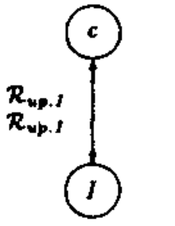

Figure 1: MBI - Ideal reasoner

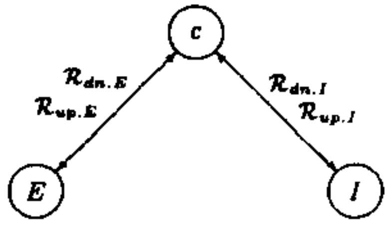

Figure 2: MBIE - Ideal and real reasoner

$C(\Omega_E, \Delta_E) \subset C(\Omega_I \cap L_E, \Delta_E)$. The second possibility is that $\Omega_E \not\subset \Omega_I$. Technically this situation is easy to construct. It is sufficient to substitute some axioms with some new "weaker" ones which are not contained in Ω_I. The issue here is whether an E so defined can still be considered the realized version of I. Intuitively one would like to have an operation of realization which amounts to limiting the capabilities of the ideal reasoner without "creating" anything really new. However consider the situation where I believes A and $A \supset B$ and E believes B. In this case it seems plausible to say that E is a realization of I.

Incompleteness in the deduction rules ($\Delta_E \prec \Delta_I$). This case is entirely dual to that for W_E. That is, we may have ignorance of some rules ($\Delta_E \subset \Delta_I$) or resource boundedness ($\Delta_E \not\subset \Delta_I$).

3 MC systems for belief

In this paper we take the computer to be logically omniscient. Even under this hypothesis, we are able to build systems whose provability relations are equivalent to those of all the systems defined in the past.

3.1 Saturated belief

We start by considering the case with an agent which is the ideal reasoner I described above. We have two contexts, one for I and one for the computer c having beliefs about I's beliefs. We take IB to be the belief predicate for I in L_c. For IB to mean (in c) belief of I we require that the following two properties hold:

- c must be complete for I: for any wff A believed by I, IB("A") is believed by c.
- c must be correct for I: for any wff of the form IB("A") believed by c, A is believed by I.

Technically, these properties are obtained via certain bridge rules (*i.e.* rules whose premises and conclusions belong to distinct contexts) which allow us to prove IB("A") in c just because A is a belief of I and, viceversa, to prove A in I just because IB("A") is a belief in c. Figure 1 gives a structural description of the resulting system, called MBI. MBI is defined as follows (MC systems are pairs "set of contexts, set of bridge rules". We write $\langle A, i\rangle$ to mean A and that $A \in L_i$. We say that A is an L_i-wff to mean that $A \in L_i$. The formal definition of deduction of an MC system is given in [Giunchiglia, 1993].)

Definition 3.1 (MBI) *Let P be a set of primitive propositional letters. An MC system $MBI = \langle\{c, I\}, \Delta_{cI}\rangle$ is such that:*

(i) $c = \langle L_c, T_c\rangle$. *$L_c$ is a propositional language whose atomic formulas are P and* IB("A") *for any formula A in L_I. L_c is closed under the propositional connectives. $T_c = C(\Omega_c, \Delta_c)$. Δ_c is correct and complete for propositional logic.*

(ii) $I = \langle L_I, T_I\rangle$. *$L_I$ is a propositional language whose atomic formulas are P. L_I is closed under the propositional connectives. $T_I = C(\Omega_I, \Delta_I)$. Δ_I is correct and complete for propositional logic.*

(iii) *Δ_{cI} is the following set of bridge rules:*

$$\frac{\langle A, I\rangle}{\langle \text{IB}(\text{“}A\text{”}), c\rangle}\ \mathcal{R}_{up.I} \qquad \frac{\langle \text{IB}(\text{“}A\text{”}), c\rangle}{\langle A, I\rangle}\ \mathcal{R}_{dn.I}$$

Restrictions: *$\mathcal{R}_{up.I}$ can be applied only if the index of every undischarged assumption $\langle A, I\rangle$ depends on, is equal to c.*

MBI was already discussed in [Giunchiglia and Serafini, 1991]. $\mathcal{R}_{up.I}$ and $\mathcal{R}_{dn.I}$ are called reflection up and reflection down respectively. $\mathcal{R}_{dn.I}$ allows to derive A in I from the fact that IB("A") has been derived in c. $\mathcal{R}_{up.I}$ has a similar reading. The restriction on $\mathcal{R}_{up.I}$ is needed to preserve the correctness of c.

3.2 Implicit and Explicit belief

One approach to the formalization of non logical omniscience is based on the distinction between explicit beliefs and implicit beliefs. According to Levesque [Levesque, 1984], while explicit beliefs are the set of effective beliefs of the agent, implicit beliefs model what the world would be like if what he believed were true. Implicit belief is used as a limit notion which gives an upper bound to what a real agent can ever believe. No matter what it explicitly believes, it will never explicitly believe more than its implicit beliefs. In this context, a natural step is to take the agent's implicit and explicit beliefs as the beliefs of an ideal reasoner *1* and of its realized version *E,* respectively. The intuition is in fact the same, that is that I/ implicit belief is the limit idealized version of *E/* explicit belief. However the emphasis is slightly changed as we do not care of how the world is but only of the subjective view of /. Implicitly we assume that the world is how / assumes it is, but this is irrelevant from a technical point of view.

Thus, if we want to model a situation where *c* can reason about the idealized and real capabilities of an agent,

we have to have two contexts I and E and two belief predicates, that is IB and EB. As we want c to be a perfect believer for I and E, we have the two reflection rules for each of the two contexts. Figure 2 gives a structural description of the resulting MC system, called *MBIE*. The formal definition follows:

Definition 3.2 (MBIE) *Let P be a set of primitive propositional letters. An MC system $MBIE = \langle\{c, I, E\}, \Delta_{cI} \cup \Delta_{cE}\rangle$ is such that:*

(i) *c is as in definition 3.1 with the atomic formulas of L_c extended to contain also EB("A"), for any formula $A \in L_I$;*

(ii) *I is as in definition 3.1;*

(iii) *Δ_{cI} is as in definition 3.1;*

(iv) *$E = \langle L_E, T_E\rangle$ with $P_E \preceq P_I$, $W_E \preceq W_I$, $\Omega_E \preceq \Omega_I$ and $\Delta_E \preceq \Delta_I$;*

(v) *Δ_{Ec} is the following set of bridge rules:*

$$\frac{\langle A,\ E\rangle}{\langle \mathrm{EB}(\text{“}A\text{”}),\ c\rangle}\ \mathcal{R}_{up.E} \qquad \frac{\langle \mathrm{EB}(\text{“}A\text{”}),\ c\rangle}{\langle A,\ E\rangle}\ \mathcal{R}_{dn.E}$$

Restrictions: *$\mathcal{R}_{up.E}$ can be applied only if the index of every undischarged assumption $\langle A,\ E\rangle$ depends on, is not equal to E; $\mathcal{R}_{dn.E}$ can be applied only if A is a L_E-wff.*

Some observations. E is defined to be the realized version of I. According to clause (i), even if a certain formula A is not an L_E-wff, EB("A") is a wff of L_c. This allows us to say in c that A is not explicitly believed ($\neg$EB("A")).

3.3 Local reasoning

Underlying all the above definitions is the intuition that a reasoner can use all its construction rules on all its basic facts. However, it is well known that human beings have difficulty in putting together all the information they possess. Human memory seems structured in frames of mind hardly communicating between them [Stalnaker, 1984). Reasoning happens locally to each frame and hardly interacts with the reasoning performed in other frames. Thus, for instance, we never think of Africa and a flat tire at the same time. The obvious way to model this situation is to associate to each agent a set of reasoners $E_i, \ldots, E_n$, one for each frame of mind, variously interconnected among themselves and with *c*. This can be done in many different ways, each modeling a different intuition. In the following we discuss one very general possibility.

As usual *c* is connected to I via reflection up and reflection down. We suppose that the various E_i have different languages and different sets of inference rules. We have a reflection up rule between *c* and each E_i. This guarantees the completeness of c. However we have only one predicate EB for all frames and no reflection down rules into the frames. This models the intuition that *c* has a correct view of what is explicitly believed but without knowing in which frame theorems are proved. EB("A") means that there is at least one frame of mind in which

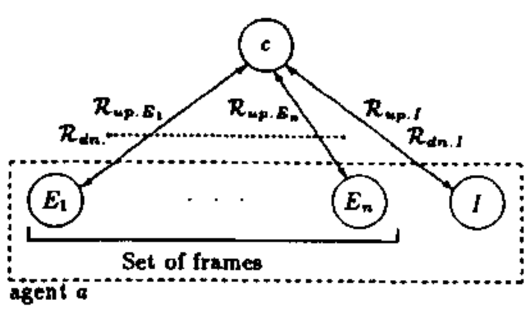

Figure 3: MBIE(n) - Local reasoning

A is true. A weak form of reflection down and therefore a weak notion of correctness for c can still be given. Suppose that we have n frames $E_1, \ldots, E_n$ of which A is a formula. Suppose that in all such frames but E_1, assuming A we derive a formula B that we do not believe. Then, as a consequence, A must be believed in E_1. This models the intuition that, still not knowing a priori in which frame a theorem is proved, c can sometimes infer this by evaluating all the possibilites. The structure of the resulting system, called MBIE(n), is given in figure 3. Its definition is as follows.

Definition 3.3 (MBIE(n)) *Let P be a set of primitive propositional letters and n a natural number. An MC system $MBIE(n) = \langle\{c, I, \{E_i\}_{1\leq i\leq n}\}, \Delta_{cI} \cup \Delta_{cE(n)}\rangle$ is such that:*

(i) *c is as in definition 3.2;*

(ii) *I is as in definition 3.2;*

(iii) *Δ_{cI} is as in definition 3.2;*

(iv) *Each $E_i = \langle L_{E_i}, T_{E_i}\rangle$ is defined the same as E in definition 3.2;*

(v) *$\Delta_{cE(n)}$ contains bridge rules $\mathcal{R}_{up.E_i}$, $(1 \leq i \leq n)$*

$$\frac{\langle A,\ E_i\rangle}{\langle \mathrm{EB}(\text{“}A\text{”}),\ c\rangle}$$

and bridge rules $\mathcal{R}_{dn.E_{k_1}}$, $(1 \leq k_1 \ldots k_m \leq n)$:

$$\frac{\langle \mathrm{EB}(\text{“}A\text{”}),\ c\rangle \quad \{\langle \neg\mathrm{EB}(\text{“}A_i\text{”}),\ c\rangle\}_{2\leq i\leq m} \left\{\begin{matrix}[\langle A,\ E_{k_i}\rangle] \\ \langle A_i,\ E_{k_i}\rangle\end{matrix}\right\}_{2\leq i\leq m}}{\langle A,\ E_{k_1}\rangle}$$

Restrictions: *$\mathcal{R}_{up.E_i}$ can be applied only if the index of every undischarged assumption $\langle A,\ E_i\rangle$ depends on, is not equal to E_i; in the rule $\mathcal{R}_{dn.E_{k_1}}$: $k_1, k_2, \ldots, k_m$ are all the indexes such that A is a wff of the corresponding frame.*

Definition 3.3 generalizes definition 3.2 of MBIE in the sense that E has been substituted with n frames.

Notice that, starting from the intuition that a reasoner can be modeled as a logical theory, we have introduced various forms of non logical omniscience and associated them to a particular structural part of E. The resulting system MBIE(n) models all of them. This process is exhaustive in the sense that all the structural subparts of E have been considered (that is its language, its theorems and E itself). If one accepts the abstract characterization of reasoners as contexts, this process is exhaustive also in the sense that it considers all the possible forms of non logical omniscience. Experimentally, so far we have failed to find a form of non logical omniscience which cannot be modeled as composition of the basic forms defined above. In particular, all the forms of lack of logical omniscience defined in the past that we are aware of can be uniformly reconstructed (see next section for some hints of how this can be done).

It can be proved that the following proposition holds ($\vdash_{MS}$ is the consequence relation defined by the MC system MS):

Proposition 3.1 *For any L_I-wff A_1, A_2 and any L_{E_i}-wff B_1, B_2:*

- (i) *if A_1 is a classical tautology then* $\vdash_{MBIE(n)} \langle IB(``A_1"), c\rangle$;
- (ii) $\vdash_{MBIE(n)} \langle IB(``A_1") \wedge IB(``A_1 \supset A_2") \supset IB(``A_2"), c\rangle$;
- (iii) *if* $\langle A_1, I\rangle \vdash_{MBIE(n)} \langle A_2, I\rangle$, *then* $\vdash_{MBIE(n)} \langle IB(``A_1") \supset IB(``A_2"), c\rangle$;
- (iv) *if* $\langle B_1, E_i\rangle \vdash_{MBIE(n)} \langle B_2, E_i\rangle$ *for any* $1 \leq i \leq n$, *such that B_1 and B_2 are L_{E_i}-wff, then* $\vdash_{MBIE(n)} \langle EB(``B_1") \supset EB(``B_2"), c\rangle$.
- (v) *if $\Omega_c = \emptyset$, then if* $\vdash_{MBIE(n)} \langle EB(``B_1"), c\rangle$ *then* $\vdash_{MBIE(n)} \langle IB(``B_1"), c\rangle$.

Item (i) says that all classical tautologies are implicitly believed. Item (ii) says that implicit belief is closed under implication. Item (iii) says that implicit belief simulates (in c) deduction in I. Item (iv) states the same property as (iii) for explicit belief. Item (v) says that explicit beliefs are implicitly believed. This is in fact a consequence of the fact that $T_{E_i} \subseteq T_I$, with $1 \leq i \leq n$. However, from the point of view of c, there is no apparent relation between E's theorems and I's. c can be made aware of this relation by adding a new set of bridge rules each connecting a frame E_i to I saying that a formula can be derived in I because of it has been derived in E_i.

Definition 3.4 (MBIE'(n)) MBIE'(n) *is an MC system obtained from* MBIE(n) *by adding, for all frames E_i the following bridge rule:*

$$\frac{\langle A, E_i\rangle}{\langle A, I\rangle}\ E_i2I$$

Adding E_i2I models the intuition that c knows that E believes a subset of I's beliefs and that explicitly uses this information to build proofs in I. Adding E_i2I does not change what is believed by I and E. However in MBIE'(n), $EB(``A") \supset IB(``A")$ is a theorem of c (while this is not the case for MBIE(n)).

Proposition 3.2 *For each L_{E_i}-wff A,* $\vdash_{MBIE'(n)} \langle EB(``A") \supset IB(``A"), c\rangle$.

4 Some important instances

It is now interesting to specialize definition 3.3 of MBIE(n) to consider MC systems where each form of lack of logical omniscience is taken in isolation. This allows us to define interesting MC systems some of which can be proved equivalent to the various systems defined in the past (for instance all those in [Fagin and Halpern, 1988; Levesque, 1984; Konolige, 1984]). In the following we study, the situation of lack of awareness. Let us consider the following MC system.

Definition 4.1 (MLL, MLL') *Let MLL (MLL') be an MC system obtained from MIEB(1) (MIEB'(1)) by imposing $P_E \prec P_I$ and $\Delta_E = \Delta_I$.*

The intuition behind MLL and MLL' is very similar to that behind Fagin and Halpern's *logic of general awareness* (called LAW from now on) [Fagin and Halpern, 1988]. In LAW an agent explicitly believes a subset of the formulas he is aware of. This is obtained by extending the modal language with a modal operator of awareness, AW. An agent explicitly believes α if and only if he is aware of α and implicitly believes it, in formulas

$$EB\alpha \equiv IB\alpha \wedge AW\alpha \tag{1}$$

(1) is in fact an axiom schema of LAW. Suitable axioms define the set of formulas of which the agents are aware of, *e.g.* there may be axioms stating that awareness is closed with respect to certain subformulas. MLL, MLL' and LAW can be proved to prove exactly the same set of implicit and explicit beliefs (by taking the obvious translation where modal operators are translated into unary predicates. For any modal formula A, we write A^+ to mean its first order translation) [2].

Theorem 4.1 *Let LAW a system for general awareness such that all the axioms involving AW but (1) are of the form* $AW\alpha_1 \wedge \ldots \wedge AW\alpha_n \supset AW\alpha$ *with $n \geq 0$. Then, for any L_c-wff A,* $\vdash_{MLL} \langle A, c\rangle$ ($\vdash_{MLL'} \langle A, c\rangle$) *if and only if* $\vdash_{LAW} A^+$.

(The condition on the axioms involving AW does not involve anything deep and could be lifted away. We have imposed it to guarantee not circular definitions of MC systems. In principle one could use AW to define L_E in terms of T_E.) However the inverse of theorem 4.1 cannot be proved, *i.e.* LAW cannot be proved equivalent to MLL nor to MLL'. Notice that in theorem 4.1 we quantify over L_c-wffs and not over the wffs in LAW. The problem is that LAW's language contains the modal operator AW which does not have any counterpart in MLL or MLL'. This can be easily solved by adding a monadic awareness predicate AW to L_c and by making sure that the formulas of the kind $AW(``\alpha")$ belong to $T_c$, subject to the restriction that $\alpha$ is a formula of $L_E$. This step makes the translation of one direction of axiom (1), that is $EB(``A") \supset IB(``A") \wedge AW(``A")$ a theorem of MLL' (but not of MLL). However this is not the case for the other direction, that is $IB(``A") \wedge AW(``A") \supset EB(``A")$. This effect can be obtained by adding MLL' the axiom $AW(``A") \equiv IB(``A") \supset EB(``A")$ and a new bridge rule, called $I2E$, which allows to derive A in E whenever A has been derived in I, subject to the restriction that A is a wff in L_E.

[2] All the modal systems considered in this paper are given semantically. In comparing a semantically presented modal system S_1 with an MC system MS_1, we say that S_1 and MS_1 are equivalent, i.e. $A_1, \ldots, A_n \models_{S_1} A$ if and only if $\langle A_1^+, c\rangle, \ldots, \langle A_n^+, c\rangle \vdash_{MS_1} \langle A^+, c\rangle$.

Definition 4.2 (MLAW) *Let MLAW be an MC system obtained from MLL' by extending L_c with the monadic predicate AW, Ω_c with the axiom* AW("A") $\equiv$ IB("A") $\supset$ EB("A") *and by adding the following bridge rule:*

$$\frac{\langle A, I\rangle}{\langle A, E\rangle}\ I2E$$

Restriction: A is an L_E-wff.

Theorem 4.2 *LAW and MLAW are equivalent.*

In order to get the equivalence result with LAW we had to extend MLL in various ways. However it should be clear that MLL (and not MLAW) satisfies all and only the minimal structural properties of agents with lack of awareness, that is the fact that $L_E \subset L_I$. The extra conditions satisfied by MLAW do not change the beliefs of I and E (theorem 4.1). Their effect is to allow c to prove more and more facts about the lack of logical omniscience of E (proposition 3.2, theorem 4.2). However, while the addition of AW to L_c and of $E_1 2I$ to the set of bridge rules of MLL has an intuitively plausible interpretation, this is not the case for $I2E$. $I2E$ says that an implicit belief which can be expressed in the language of L_E is also an explicit belief. As a consequence, it allows derivations (in general not possible otherwise) where explicit beliefs are first proved in I and then exported to E. However we do not find derivations of this kind intuitively plausible as we think that E's deductive capabilities should not exploit I's. Finally, the axiom added to MLL' to obtain MLAW seems to exist mainly for technical reasons.

As a general observation, the advantages over a single theory approach should now be clear. In MC systems, localization (of the language, of the axioms, of the deductive machinery) is intrinsic in the logic, namely in the fact that we have distinct contexts. The properties of an agent can be directly imposed on the contexts representing it and not, as it is the case in a single theory approach, in some global theory of the world (playing the role of c). This allows for easier (to do and to understand) formalizations where each agent is locally defined independently of anything else.

5 Example

Let us consider MBIE(2), that is the MC system that formalizes the beliefs of an agent with two frames of mind. The following wff is provable in MBIE(2).

$$\langle(\text{EB}(\text{“}A\text{”}) \wedge \text{EB}(\text{“}B\text{”}) \wedge \text{EB}(\text{“}C\text{”})) \supset (\text{EB}(\text{“}A \wedge B\text{”}) \vee \text{EB}(\text{“}A \wedge C\text{”}) \vee \text{EB}(\text{“}B \wedge C\text{”})), c\rangle$$

Intuitively it says that if A, B and C are explicit beliefs, then there is a frame where one among $A \wedge B$, $A \wedge C$ and $B \wedge C$ is provable. Notice that, as a consequence, at least two among A, B and C are formulas of this frame. The proof is given in figure 5. Notice that inside each context we use Natural Deduction and some deciders, *e.g.* the decider for propositional logic TAUT [Giunchiglia, 1992]. The proof in figure 5 can be summarized as follows. We proceed by contradiction and assume in context c the hypothesis , *i.e.* EB("A") $\wedge$ EB("B") $\wedge$ EB("C") (line 4) and the negation of the conclusion, *i.e.* $\neg$EB("$A \wedge B$"), $\neg$EB("$A \wedge C$") and $\neg$EB("$B \wedge C$") (line 5). Assuming B in E_1 (line 2), from the fact that EB("A") and $\neg$EB("$A \wedge B$") are derived in c (lines 6, 7) and from the fact that there are only two frames, we conclude that A is derivable in E_2. (line 8). The hypothesis EB("C") (line 11) and the fact that there are two frames implies that C must be derivable in E_1 or in E_2. In the first case we have $B \wedge C$ in E_1 (line 14) which implies EB("$B \wedge C$") in c but this leads us to a contradiction (line 16). Exactly the same argument can be given for E_2 and $A \wedge C$ (lines 10, 12). This allows us to infer A in (line 17) and then $A \wedge B$ in E_2 (line 29). Reflecting up $A \wedge B$ (line 30) allows us to derive the goal by propositional reasoning in c (line 34).

A first observation is that the agents' contexts are used to prove theorems which are then reflected up *(e.g.* line 30). This is a practical consequence of the fact that agents are not simulated in c but are rather "used" to extract what it is proved in their contexts. In this perspective reflection down is used as a means to impose a set of facts into a context *(e.g.* lines 8, 13, 17). Notice that all the applications of reflection down are above all the applications of reflection up. This is in fact the general way of performing reasoning with MC systems for belief: (i) first some conjectures are made in c; (ii) then these conjectures are "imposed" on the reasoners (via reflection down); (iii) then the reasoning capabilities of the reasoners are used io compute the consequences of the conjectures; (iv) finally the result is exported back to *c* via reflection up. Some amount of propositional reasoiling can be performed in *c* before step (i) and after step (iv).

6 Conclusion

In this paper we have proposed a new approach to the formalization of reasoning about belief where agents are modeled as sets of interacting contexts. In this framework the properties of each agent can be directly formalized by imposing local suitable conditions on the theories modeling it (that is, on the language, on the axioms, on the deductive machinery, on the bridge rules).

References

[Dinsmore, 199l] J. Dinsmore. *Partitioned Representations.* Kluwer Academic Publisher, 1991.

[Fagin and Ilalpern, 1988] R. Fagin and J.Y. Halpern Belief, awareness, and limited reasoning. *Artif. Intell* , 34:39-76, 1988.

[Giunchiglia and Serafini, 1991] F. Giunchiglia and L. Serafini. Multilanguage hierarchical logics (or: how we can do without modal logics). To *appear in: Artificial Intelligence journal,* 1991. Short version appeared in Proc. AI'92, Canadian Conference of Artificial Intelligence.

[Giunchiglia and Weyhrauch, 1988] F. Giunchiglia and R.W. Weyhrauch. A multi-context monotonic axiomatizafion of inessential non-monotonicity. In D. Nardi

Switch To Context E_1			
(1)	A	(1)	by assumption;
(2)	B	(2)	by assumption;
(3)	$A \wedge B$	(1)(2)	from (1) and (2) by $\wedge I_{E_1}$;
Switch To Context c			
(4)	$(EB(``A") \wedge EB(``B") \wedge EB(``C"))$	(4)	by assumption;
(5)	$\neg(EB(``A \wedge B") \vee EB(``A \wedge C") \vee EB(``B \wedge C"))$	(5)	by assumption;
(6)	$\neg EB(``A \wedge B")$	(5)	from (5) by TAUT;
(7)	$EB(``A")$	(4)	from (4) by $\wedge E_c$;
Switch To Context E_2			
(8)	A	(2)(4)(5)	from (3), (6) and (7) by $\mathcal{R}_{dn.E_2}$;
(9)	C	(9)	by assumption;
(10)	$A \wedge C$	(2)(4)(5)(9)	from (8) and (9) by $\wedge I_{E_2}$;
Switch To Context c			
(11)	$EB(``C")$	(4)	from (4) by $\wedge E_c$;
(12)	$\neg EB(``A \wedge C")$	(5)	from (5) by TAUT;
Switch To Context E_1			
(13)	C	(2)(4)(5)	from (10), (11) and (12) by $\mathcal{R}_{dn.E_1}$;
(14)	$B \wedge C$	(2)(4)(5)	from (13) and (2) by $\wedge I_{E_1}$;
Switch To Context c			
(15)	$EB(``B")$	(4)	from (4) by $\wedge E_c$;
(16)	$\neg EB(``B \wedge C")$	(5)	from (5) by TAUT;
Switch To Context E_2			
(17)	B	(4)(5)	from (14), (15) and (16) by $\mathcal{R}_{dn.E_2}$;
⋮			
(28)	A	(4)(5)	write A for B and B for A in (8)-(16);
(29)	$A \wedge B$	(4)(5)	from (17) and (28) by $\wedge I_{E_2}$;
Switch To Context c			
(30)	$EB(``A \wedge B")$	(4)(5)	from (29) by $\mathcal{R}_{up.E_2}$;
(31)	$EB(``A \wedge B") \vee EB(``A \wedge C") \vee EB(``B \wedge C")$	(4)(5)	from (30) by $\vee I_c$;
(32)	$\bot$	(4)(5)	from (31) and (5) by $\supset E_c$;
(33)	$EB(``A \wedge B") \vee EB(``A \wedge C") \vee EB(``B \wedge C")$	(4)	from (32) by $\bot_c$;
(34)	$(EB(``A") \wedge EB(``B") \wedge EB(``C")) \supset (EB(``A \wedge B") \vee EB(``A \wedge C") \vee EB(``B \wedge C"))$		from (33) by $\supset I_c$;

Figure 4: A proof in MBIF(2)

and P. Maes, editors, *Meta-level architectures and Reflection,* pages 271-285. North Holland, 1988.

[Ciunchiglia *et al.*, 1992] F. Ciunchiglia, L. Serafini, and A. Simpson. Hierarchical meta-logics: intuitions, proof theory and semantics. In *Proc. META-92, Workshop on Metaprogramming in Logic,* number 649 in LNCS, pages 235-249, Uppsala, Sweden, 1992. Springer Verlag.

[Ciunchiglia, 1991] F. Giiunchiglia. Multilanguage systems. In *Proc. of AAAI Spring Symposium on Logical Formalizations of Commonscnse Reasoning,* 1991.

[Ciunchiglia, 1992] F. Ciunchiglia. The GETFOL Manual - GETFOL version 1. Technical Report 9204-01, DIST - University of Cenova, Genoa, Italy, 1992.

[Ciunchiglia, 1993] F. Ciunchiglia. Contextual reasoning. *Epistemologia, special issue on I Linguaggi e le Macchine,* XVI.345 364, 1993. Short version in Proc. IJCAI'93 Workshop on Using Knowledge in its Context, Chambery, France, 1993.

[Halpern and Moses, 1985] J.Y. Halpern and Y. Moses. A guide to the modal logics of knowledge and belief: preliminary draft. In *Proc. of the 9th International Joint Conference on Artificial Intelligence*, pages 480-490, Los Angeles, CA, 1985. Morgan Kaufmann Publ. Inc.

[Konolige, 1982] K. Konolige. A First-Order Formalisation of Knowledge and Action for a Multi-Agent Planning System. In *Machine Intelligence 10,* 1982

[Konolige, 1984] K. Konolige. *A deduction model of belief and its logics.* PhD thesis, Stanford University CA, 1984.

[Levesque, 1984] II. Levesque. A logic of implicit and explicit belief. In *AAAI-84* pages 198-202, 1984.

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

[Weyhranch, 1980] R.W. Wryhranch. Prolegomena to a Theory of Mechanized Formal Reasoning. *Artif. Intell.,* 13(1):133-176, 1980.

[Wilks and Biem, 1979] Y. Wilks and J. Biem. Speech acts and multiple environments. In *Proc. of the 6th International Joint Conference on Artificial Intelligence,* 1979.