

Formalizing Belief Reports the Approach and a Case Study*

Massimo Benerecetti^{1,2}, Paolo Bouquet^{2,3}, and Chiara Ghidini²

¹ DISA – University of Trento, Via Inama 5,
I-38100 Trento, Italy

² DIST – University of Genova, Via Opera Pia 13,
I-16100 Genova, Italy

³ Department of Philosophy – University of Genova, Via Balbi 4,
I-16100 Genova, Italy

bene@cs.unitn.it bouquet@cs.unitn.it ghidini@cs.unitn.it

Abstract. One of the most interesting puzzles in formalizing belief contexts is the fact that many belief reports can be given both an opaque and a transparent readings. A traditional explanation is that the two readings are related to the failure and success of the *principle of substitutivity* respectively, and this in turn is explained with the *de re/de dicto* distinction. We propose an alternative analysis, based on the idea that another agent's beliefs can just be quoted (preserving opacity) or translated into the reporter's language (allowing for transparency). We show that MultiContext systems allow for the formalization of these two phenomena at the same time, thanks to their multi-language feature.

1 Beliefs and Substitutivity

A very important part of our intelligence is the ability to reason about our own and other people's beliefs. We have beliefs about our beliefs; we learn from other people's beliefs; we accept (don't accept) other people's beliefs among our beliefs; we share our beliefs with other people; we report other people's beliefs; we infer new information by combining different people's beliefs; we make decisions based on conjectures about other people's beliefs; many communication conventions are based on what we believe other people believe; and so on.

Despite the pervasiveness of the notion of belief and the apparent easiness for humans to deal with it, representing beliefs and formalizing reasoning with and about beliefs raises very difficult problems. From a logical point of view, a well-known problem is that a very general and intuitive principle, the *principle of substitutivity*, fails in its unrestricted version¹. Intuitively, the *unrestricted principle of substitutivity* (UPS) states that whenever two terms (two sentences)

* We thank the Mechanized Reasoning Group at DISA (Trento), ITC-IRST (Trento) and DIST (Genoa). This work is part of the MRG project *Distributed Representations and Systems* (<http://www.cs.unitn.it/~mrg/distributed-intelligence/>).

¹ This fact was first noticed in [3].

have the same meaning², they can be substituted one for the other in any sentence *salva veritate*. Let Φ be a sentence and α, α' two terms (sentences). If $\alpha = \alpha'$ ($\alpha \equiv \alpha'$), then $\Phi \equiv \Phi[\alpha/\alpha']$, where $\Phi[\alpha/\alpha']$ is the result of replacing some occurrences of α with α' . But UPS is not sound (validity preserving) when applied in belief contexts (namely within the scope of an operator of the form *X-believes-that* [...], where [...] stands for any sentence).

Consider the following instance of UPS:

<i>Mr. A believes that the president of the local football team is a corruptor</i>
<i>The president of the local football team is Mr. C</i>
<i>Mr. A believes that Mr. C is a corruptor</i>

It is easy to see that the conclusion does not logically follow from the premisses. Indeed, even though Mr. *C* is the president of the local football team, Mr. *A* could not be aware of this fact. Given the two premisses, we have no means to conclude whether Mr. *A* believes or not that the president of the local football team is Mr. *C*. Hence it is possible to think of a model in which the premisses are true and the conclusion is false. The traditional solution is to introduce a *restricted principle of substitutivity* (RPS): whenever two terms (two sentences) that occur within the scope of the same belief operator have the same meaning, they can be substituted one for the other in any sentence that occurs in the same context *salva veritate*. More formally, let Φ be a sentence, α a term (sentence), and *Bel* a belief operator. If $\text{Bel}(\alpha = \alpha')$ [$\text{Bel}(\alpha \equiv \alpha')$], then $\text{Bel}(\Phi) \equiv \text{Bel}(\Phi[\alpha/\alpha'])$. In the example above, RPS can be applied only if the second premiss is replaced by the fact that Mr. *A* believes that the president of the local football team is Mr. *C*. In this case, the conclusion holds.

However, there are cases in which this weaker form of substitutivity is not sufficient in order to model common forms of reasoning about beliefs. A simple example is the way people report other people's beliefs. Following the analysis of [1], we show that many belief reports can be given two readings, called *opaque* and *transparent* respectively, and that the second requires the application of a form of substitutivity stronger than RPS. We argue that both readings are intuitively plausible, and therefore an adequate formalization of beliefs should allow us to model both of them, and not just to eliminate one. To this end, we present and discuss in detail a motivating scenario, which we use also as a case study; then we present a multi-context framework for belief contexts and we show that it allows us to model both the opaque and the transparent readings of belief reports.

2 Opacity and Transparency in Belief Reports

The scenario we consider is a slight modification (and translation) of an example from [1]:

² The word 'meaning' is used here informally.

You know that Mr. *A* believes that the president of the local football team is Mr. *M* and you know that Mr. *B* believes that the president is Mr. *C*. You know also that Mr. *B* knows that *A* believes that the president of the local football team is Mr. *M*. Actually, Mr. *B* is right, and you know that. Now, *B* tells you: “Mr. *A* believes that the president of the local football team is a corruptor”. If you want to know whom the sentence is about (besides Mr. *A*), how will you interpret the sentence?

Suppose that the problem is posed to a computer program. The program is a little puzzled, since the question has two possible answers:

- (i) Mr. *A*'s belief is referred to Mr. *M* (since Mr. *A* is the subject of the belief), and thus the program concludes that Mr. *A* believes that Mr. *M* is a corruptor;
- (ii) Mr. *A*'s belief is referred to Mr. *C* (since it is Mr. *B* who is speaking), and thus the program concludes that Mr. *A* believes that Mr. *C* is a corruptor (even though Mr. *A* does not identify Mr. *C* as the president of the local football team).

The first is an *opaque reading* of the belief report, and requires an application of RPS: since the two beliefs that the president of the local football team is a corruptor and that the president of the local football team is Mr. *M* occur in the same belief context (i.e. Mr. *A*'s beliefs), the program can apply RPS and conclude that Mr. *A* believes that Mr. *M* is a corruptor. The second is a *transparent reading*, and requires a stronger principle than RPS. Indeed, the program must be able to reason about the fact that Mr. *B* might have used the description ‘the president of the local football team’ to mean Mr. *C* (the person Mr. *B* himself believes to be the president) and not to mean Mr. *M* (the person Mr. *A* believes to be the president). The question is whether the transparent reading is (intuitively) legitimate. The first reaction is that it is somewhat unnatural. Since Mr. *A* would not agree with the conclusion that Mr. *C* is a corruptor, the transparent reading is not “faithful” to Mr. *A*'s belief and therefore – one might conclude – incorrect. However, as it is convincingly argued in [1], this conclusion is based on the idea that only Mr. *A*'s beliefs are relevant, and completely disregards the rôle of the reporter. Let us imagine, for example, the following scenario: Mr. *B* is told by Mr. *A*: “Mr. *C* is a corruptor”; since Mr. *B* believes that Mr. *C* is the president of the local football team, reporting Mr. *A*'s belief using the description ‘the president of the local football team’ seems perfectly acceptable (even though Mr. *B* knows that Mr. *A* would not agree with this report).

Traditionally, the two readings are explained with the *de re/de dicto* distinction (e.g. in [11]): the opaque reading is associated with a *de dicto* belief and the transparent reading with a *de re* belief. However, this does not work in the proposed example. Let us consider another version of the story. Mr. *B* is told by Mr. *A*: “The president of the local football team is a corruptor”, and Mr. *B* has good reasons to think that Mr. *A* is not really referring to Mr. *M* (perhaps Mr. *A* is reporting a radio news on the president of the local football team). So,

even though Mr. *B* knows that for Mr. *A* this entails the false belief that Mr. *M* is a corruptor, Mr. *B* can report Mr. *A*'s belief, being sure that the program will understand the description 'the president of the local football team' in the right way (namely, as referring to Mr. *C*). In this case Mr. *A* has a *de dicto* belief about the president of the local football team, but Mr. *B*'s report is based on a transparent reading. So the *de re/de dicto* distinction is orthogonal to the opacity/transparency problem.

Our analysis emphasizes the conceptual spaces of all agents involved in the scenario, in particular the conceptual space of the reporter (in the example, Mr. *B*). Mr. *B* can assume two attitudes towards Mr. *A*'s reported belief. The first is to *quote* Mr. *A*'s words (something like: He told me that "[...]"). The second is to *translate* the content of Mr. *A*'s belief in his own words (something like: I would express his belief as [...]). If the program is to reason on belief reports, it must be able to ascribe both attitudes to the reporter. In particular, ascribing the first attitude will result in an opaque reading of a belief report and ascribing the second will result in a transparent reading. So let us look again at the example from this perspective. Mr. *B* tells the program that Mr. *A* believes that the president of the local football team is a corruptor. If the program ascribes to Mr. *B* the "quotation" attitude, it will reason as follows: since Mr. *B* is reporting the exact words of Mr. *A*, and since Mr. *A* believes that the president of the local football team is Mr. *M*, the content of the reported belief is that Mr. *A* believes that Mr. *M* is a corruptor (opaque reading). If the program ascribes to Mr. *B* the "translation" attitude, it will reason as follows: since Mr. *B* believes that the president of the local football team is Mr. *C*, and since the description 'the president of the local football team' is to be read in Mr. *B*'s sense, the reported belief is that Mr. *C* is a corruptor (transparent reading).

In order to formalize these intuitions, we introduce the notion of *view*. A view is a representation of a collection of beliefs that a reasoner (in our example, the program) ascribes to an agent (including itself) under a given perspective. Possible perspectives are: the beliefs that the program ascribes to itself (e.g. that Mr. *B* believes that Mr. *A* believes that the president of the local football team is a corruptor); the beliefs that the program ascribes to Mr. *B* (e.g. that Mr. *A* believes that the president of the local football team is a corruptor); the beliefs that the program ascribes to Mr. *B* about Mr. *A* (e.g. that the president of the local football team is a corruptor). As a convention, we use the Greek letter ϵ for the view containing the beliefs that the program ascribes to itself, and bold letters for labelling any other view. For instance $\epsilon\mathbf{B}$ is the view containing the beliefs that the program ascribes to Mr. *B* from its perspective, and $\epsilon\mathbf{BA}$ is the view containing the beliefs that the program ascribes to Mr. *A* from Mr. *B*'s perspective³. The views that the program can build can be organized in a structure like that presented in Figure 1⁴. Each circle represents a view; some

³ Since in our example no confusion can arise, from now on we will omit the prefix ϵ .

⁴ This structure can be easily generalized to n agents. For a more detailed description of the structure, a good reference is [2], where views are used to solve the a well-known

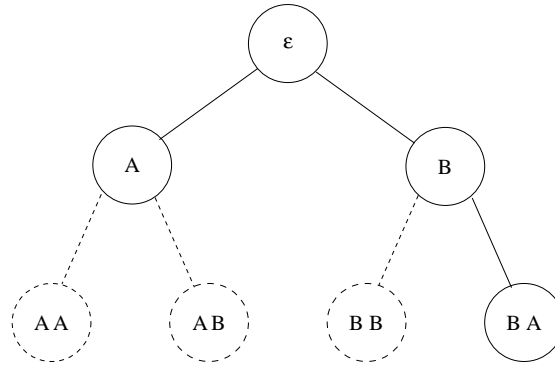


Fig. 1. The structure of views

circles are dashed because we will not use them in the formalization of the case study.

As usual, the propositional content of a belief is represented as a sentence of some language. In most traditional approaches, this language is unique and beliefs are represented as sentences of this language. Of course, a sentence can be believed by an agent and not believed by another, but it expresses one and the same proposition regardless of the agent which believes (does not believe) it. However, our case study shows that different agents may mean different propositions with the “same” sentence. This is the case of the sentence “the president of the local football team is a corruptor”: for Mr. *A*, it denotes the proposition that Mr. *M* is a corruptor; for Mr. *B*, the proposition that Mr. *C* is a corruptor. Since the definite description ‘the president of the local football team’ is such that it can refer only to a single person, the two propositions entail a contradiction (unless we can prove that Mr. *M* is the same person as Mr. *C*).

Our formalization is based on an idea proposed in [6]: there is not a single language that the program uses to describe its views about other agents’ beliefs; instead, *a distinct language is associated with each view, and the interpretation of such a language is local to the view it is associated with*. In this way, we can distinguish the sentence “the president of the local football team is a corruptor” in **B** and the sentence “the president of the local football team is a corruptor” in **BA**. Even though syntactically they “look the same”, they do not denote the same proposition, and therefore are true in two different (and in principle independent) sets of models. This property is called *locality*, and is a key point in the formalization of the intrinsic ambiguity of belief reports.

The contents of different views are obviously related. For instance, it is reasonable to assume that there is a relation between the fact that the program puts into **B** the belief that Mr. *A* believes that the president of the local football team

puzzle involving reasoning about belief and ignorance, namely the Three-Wise-Men problem.

is a corruptor, and into **BA** the belief that the president of the local football team is a corruptor. Any relation between sets of facts belonging to different views is called a *compatibility relation*. A very intuitive relation of compatibility between views is the following: if a sentence of the form ϕ belongs to **BA**, then a sentence of the form “*Mr. A believes that ϕ* ” belongs to **B** (where *believes* is a belief operator). In this case, we say that *B* is a correct observer; if the relation holds in the other direction as well (i.e. if a sentence of the form “*Mr. A believes that ϕ* ” belongs to **B**, then a sentence of the form ϕ belongs to **BA**), we say that *B* is a correct and complete observer [5]. But it is very important to realize that correct and complete observers are not the only possible kind. In fact, for the transparent reading we need a different kind of observer, defined by the following compatibility relation: if **BA** contains a sentence ϕ and **B** contains an equality (equivalence) of the form $\alpha = \alpha'$ ($\alpha \equiv \alpha'$), then a sentence of the form “*A believes^T that $\phi[\alpha/\alpha']$* ” belongs to **B**, (where *believes^T* is a belief operator distinct from *believes*). In the remain of the paper, we show – both proof-theoretically and model-theoretically – that these two compatibility relations are basically the relations between views that we need in order to model opacity and transparency in belief reports.

3 Belief Contexts in MultiContext Systems

In order to model the properties of locality and compatibility as discussed in the previous section, we formalize belief reports in the framework of *MultiContext systems* (MC systems). In this section we review only those aspects of MC systems that are needed in order to present our formalization of belief reports. The interested reader may refer to the bibliography for a more complete presentation⁵.

Formally, given a set I of indices, a MC system is a pair $\langle \{C_i\}_{i \in I}, BR_I \rangle$, where $\{C_i\}_{i \in I}$ is a collection of *contexts* and BR_I is a set of *bridge rules*. Each context is defined as an axiomatic formal system, i.e. a triple $\langle L_i, \Omega_i, \Delta_i \rangle$, where L_i is the language of C_i , $\Omega_i \subseteq L_i$ is the set of axioms of C_i and Δ_i is the set of (local) inference rules of C_i in Natural Deduction style [13]. When no confusion can arise, we will refer to contexts using their indices (i.e., we will refer to context C_i as i). Notationally we write $i : \phi$ to mean that ϕ is a formula of i . A bridge rule is defined as an inference rule with premisses and conclusion in different contexts. For instance, the bridge rule

$$\frac{1 : \phi_1}{2 : \phi_2} \tag{1}$$

says that the formula ϕ_2 is derivable in context 2 from the fact that the formula ϕ_1 is derivable in context 1.

⁵ MC systems were formally defined in [6] and in [9] (with the name of *MultiLanguage systems*). The semantics of MC systems, called *Local Model Semantics*, was presented in [8, 4].

In [10, 5, 2], it is shown how belief contexts can be formalized using MC systems. Our work is built on top of this work. The basic ideas are the following. We imagine that the computer program ϵ is able to handle a set I of views representing the collections of beliefs that it ascribes to Mr. A and Mr. B (the set of views of Figure 1). Formally, each view i is thought of as a context C_i , with its own language, set of axioms and local inference rules (locality). Compatibility among different views is modelled by defining a suitable set of bridge rules.

In Local Model Semantics (LMS), locality is modelled by associating a set of models – in our case, first order models – and a satisfiability relation (written as \models_c) – in our case, first order (classical) satisfiability – to each context. A model m of the language L_i which satisfies the set of axioms Ω_i is called a *local model* of the context C_i . Compatibility is modelled as a relation over sets of local models of different contexts. Let \mathbf{c}_i be a set of local models of the context C_i . By $\mathbf{c}_i \models \phi$ we mean that for every $m \in \mathbf{c}_i$, $m \models_c \phi$. The sequence $\mathbf{c} = \langle \mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_i, \dots \rangle$ is called a *compatibility sequence* (for $\{C_i\}_{i \in I}$). A *compatibility relation* is a set of compatibility sequences. For instance, in the case $I = \{1, 2\}$, the compatibility relation

$$\{ \langle \mathbf{c}_1, \mathbf{c}_2 \rangle \mid \text{if } \mathbf{c}_1 \models \phi_1, \text{ then } \mathbf{c}_2 \models \phi_2 \} \quad (2)$$

is the semantic counterpart of the bridge rule (1). A *model* for an MC system $\langle \{C_i\}_{i \in I}, BR_I \rangle$ is a compatibility relation \mathbf{C} defined over sets of (local) models of $\{C_i\}_{i \in I}$, modelling the “links” formalized by the bridge rules belonging to BR_I .

In the next two sections we show how to use MC systems to formalize both the opaque and the transparent reading of the belief report “Mr. A believes that the president of the local football team is a corruptor” (our case study). In order to do this, (i) we define the set of contexts (each with its set of local models) representing the views of the example; (ii) we define the bridge rules (compatibility relations) representing the fact that the computer program may perform an opaque or a transparent reading of the belief of Mr. A as reported by Mr. B .

4 Representing Opacity

Following [2], we first introduce the class of languages for the program’s views. Let L be a first order language containing two constants C and M (Mr. C and Mr. M respectively), a unary predicate $Corr$ (to be a corruptor), and a definite description ιP^6 (the president of the local football team). Let $I_0 = \{A, B\}$ be the set of the agents’ names in the scenario, and $I = I_0^*$ the set of finite (possibly empty) sequences of elements in I_0 . Let $\epsilon \in I$ denote the empty sequence. A *MultiView language* (MV language) is defined as a class of first order languages $\{L_i\}_{i \in I}$, where

- (i) $L \subseteq L_i$ for each $i \in I$;

⁶ We use the notation of [14] for definite descriptions.

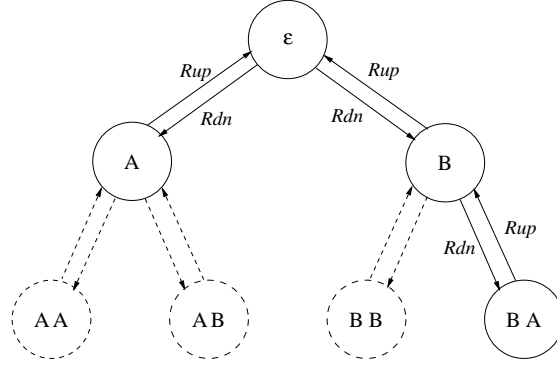


Fig. 2. The MV system

- (ii) for every agent name $a \in I_0$, if $\phi \in L_{ia}$ then the term “ ϕ ” belongs to L_i (where “ ϕ ” is the “name” of the formula ϕ);
- (iii) for every agent name $a \in I_0$, Bel_a and Bel_a^T are unary predicate symbols of L_i (where Bel_a is the predicate used by i to mean “ a told me that [...]” and Bel_a^T is the predicate used by i to mean “I would express a ’s beliefs as [...]”).

We say that an index $i \in I$ is “greater than” an index $j \in I$ if and only if j is a strict prefix of i (i.e. $i = jk$ for some $k \neq \epsilon$).

4.1 Formalizing Opacity with MC Systems

The formal system allowing for the opaque reading in our case study is defined as follows:

Definition 1. An opaque system $MV = \langle \{C_i\}_{i \in I}, BR_I \rangle$ is an MC system such that:

- I_0 and I are the sets defined above;
- for every $C_i \in \{C_i\}_{i \in I}$:
 - L_i is the language of index i in the class of MV languages defined above;
 - Ω_ϵ contains the formulae: $Bel_B(“\iota P = C”)$, $Bel_B(“Bel_A(“\iota P = M”)$), $Bel_B(“Bel_A(“Corr(\iota P)”)$);
 - $\Omega_i = \emptyset$, for any $i \neq \epsilon$;
 - Δ_i is the set of inference rules of Natural Deduction for first order logics plus a sound and complete set of rules for equality;
- the set BR_I contains, for any index $ia \in I$, the bridge rules of the form:

$$\frac{ia : \phi}{i : Bel_a(“\phi”)} R_{up} \qquad \frac{i : Bel_a(“\phi”)}{ia : \phi} R_{dn}$$

where R_{up} is applicable only if ϕ is a closed formula and no formula the premiss $ia : \phi$ depends upon has index greater than i .

Figure 2 depicts the MultiContext architecture of MV : circles correspond to contexts and labeled arrows to bridge rules among contexts. Context ϵ represents the view of the program itself, while each context ia represents the view of agent a from the point of view i .

Figure 3 shows the MultiContext proof, reported in a Natural Deduction style, of the formula $Bel_B("Bel_A("Corr(M)")")$ in the context ϵ of MV starting from the formulae in Δ_ϵ . Deduction local to a context i is surrounded by a box

$$\begin{array}{c}
\frac{\frac{\frac{\frac{\boxed{Bel_B("Bel_A(" \iota P = M ")")}_\epsilon}{\boxed{Bel_A(" \iota P = M ")}_B} R_{dn}}{\boxed{Bel_B("Bel_A("Corr(\iota P)")")}_\epsilon} R_{dn}}{\boxed{Bel_A("Corr(\iota P)")}_B} R_{dn}}{\frac{\frac{\frac{\boxed{\iota P = M}}{\boxed{Corr(M)}}}{\boxed{Corr(\iota P)}} Subst}{\boxed{Corr(M)}}_{BA} R_{up}} R_{up}}{\boxed{Bel_A("Corr(M)")}_B} R_{up}}{\boxed{Bel_B("Bel_A("Corr(M)")")}_\epsilon} R_{up}
\end{array}$$

Fig. 3. Reasoning in MV

labeled i . Bridge rules connect local deductions performed in different contexts, their premisses and conclusions being the last formulae of the boxes above and the first formulae in the boxes below them, respectively. The last formula of the deduction corresponds to the opaque reading of Mr. B 's report of Mr. A belief $Bel_B("Bel_A("Corr(\iota P)")")$. The rule $Subst \in \Delta_i$ is the standard rule of substitutivity and allows for the multi-context version of the RPS principle.

4.2 Modeling Opacity with LMS

We will now give a model-theoretic account of opacity by defining the notion of opaque model for MV .

Definition 2. A model \mathbf{C} , closed under containment, for MV is an opaque model if for every $\mathbf{c} \in \mathbf{C}$ and for every $i \in I$ and $a \in I_0$:

$$\text{if } \mathbf{c}_i \models Bel_a(" \phi ") \text{ then } \mathbf{c}_{ia} \models \phi \quad (3)$$

$$\text{if for all } \mathbf{c}' \in \mathbf{C}, \mathbf{c}'_i \subseteq \mathbf{c}_i \text{ implies } \mathbf{c}'_{ia} \models \phi, \text{ then } \mathbf{c}_i \models Bel_a(" \phi ") \quad (4)$$

Condition (3) says that i is a correct observer because if $Bel_a(" \phi ")$ is satisfied by \mathbf{c}_i then ϕ is satisfied by \mathbf{c}_{ia} . Condition (4) says that i is a complete observer. Indeed \mathbf{c}_i satisfies all the formulae $Bel_a(" \phi ")$ such that ϕ is satisfied by all the

models \mathbf{c}'_{ia} related (via compatibility relation) with (a subset of) \mathbf{c}_i , i.e. such that ϕ is believed by ia , the mental image that i has of a . Notice that (3) is “weaker” than (4). This fact reflects the difference between R_{dn} and R_{up} , namely the fact that R_{up} is a “restricted” bridge rule. This means that all of i ’s beliefs of the form $Bel_a(\phi)$ are correct even if they follow from assumptions (and not only theorems) in i . In other words, Condition (3) states the soundness of hypothetical reasoning about beliefs in an opaque model.

Now we are ready to show how ϵ may use (3) and (4) to perform the opaque reading. Let us consider an opaque model \mathbf{C} for MV . The initial knowledge of ϵ is

$$\boxed{\begin{array}{l} \epsilon \quad Bel_B(\iota P = C) \\ \quad \quad Bel_B(\text{“}Bel_A(\iota P = M)\text{”}) \\ \quad \quad Bel_B(\text{“}Bel_A(\text{“}Corr(\iota P)\text{”})\text{”}) \end{array}} \quad (5)$$

This means that any set of local models of ϵ must satisfy the formulae in (5). Constraint (3) between contexts ϵ and \mathbf{B} tells us that every compatibility sequence $\langle \mathbf{c}_\epsilon, \mathbf{c}_A, \mathbf{c}_B, \mathbf{c}_{AA}, \mathbf{c}_{AB}, \mathbf{c}_{BA}, \mathbf{c}_{BB} \dots \rangle$ is such that \mathbf{c}_B must satisfy

$$\boxed{\begin{array}{l} B \quad \iota P = C \\ \quad \quad Bel_A(\iota P = M) \\ \quad \quad Bel_A(\text{“}Corr(\iota P)\text{”}) \end{array}} \quad (6)$$

Constraint (3) between contexts \mathbf{B} and \mathbf{BA} tells us that every compatibility sequence $\langle \mathbf{c}_\epsilon, \mathbf{c}_A, \mathbf{c}_B, \mathbf{c}_{AA}, \mathbf{c}_{AB}, \mathbf{c}_{BA}, \mathbf{c}_{BB} \dots \rangle$ is such that \mathbf{c}_{BA} must satisfy

$$\boxed{\begin{array}{l} BA \quad \iota P = M \\ \quad \quad Corr(\iota P) \end{array}} \quad (7)$$

From the definition of local semantics as first order semantics, every \mathbf{c}_{BA} satisfies also the following consequence of the equality axioms:

$$\boxed{BA \quad Corr(M)} \quad (8)$$

Finally, constraint (4) between contexts \mathbf{B} and \mathbf{BA} tells us that every compatibility sequence $\langle \mathbf{c}_\epsilon, \mathbf{c}_A, \mathbf{c}_B, \mathbf{c}_{AA}, \mathbf{c}_{AB}, \mathbf{c}_{BA}, \mathbf{c}_{BB} \dots \rangle$ is such that \mathbf{c}_B must satisfy

$$\boxed{B \quad Bel_A(\text{“}Corr(M)\text{”})} \quad (9)$$

and constraint (4) between contexts ϵ and \mathbf{B} tells us that every compatibility sequence $\langle \mathbf{c}_\epsilon, \mathbf{c}_A, \mathbf{c}_B, \mathbf{c}_{AA}, \mathbf{c}_{AB}, \mathbf{c}_{BA}, \mathbf{c}_{BB} \dots \rangle$ is such that \mathbf{c}_ϵ must satisfy

$$\boxed{\epsilon \quad Bel_B(\text{“}Bel_A(\text{“}Corr(M)\text{”})\text{”})} \quad (10)$$

that is \mathbf{C} satisfies an opaque reading of Mr. B ’s report of Mr. A belief.

Steps (5)–(10) are the model-theoretic counterpart of the proof of figure 3, where the application of constraints (3) and (4) corresponds to the application of

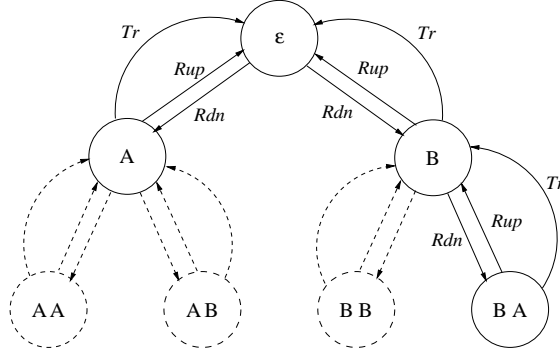


Fig. 4. The MV^T system

the bridge rules R_{dn} and R_{up} respectively. The proof that MV systems are sound and complete with respect to the class of opaque models is a straightforward generalization of the soundness and completeness theorem for the single agent case in [7].

5 Representing transparency

We are interested in modelling both the opaque and the transparent readings in a single system. Therefore, we build a formalization of transparent readings on top of the formalization of opaque readings, obtaining a formal system in which both the readings are possible.

5.1 Formalizing Transparency with MC Systems

The formal system allowing for both readings in our case study is defined as follows:

Definition 3. A transparent system $MV^T = \langle \{C_i\}_{i \in I}, BR_I \rangle$ is an opaque system containing, for any $ia \in I$, the additional bridge rules:

$$\frac{ia : \phi}{i : (\tau = \sigma) \supset Bel_a^T(\phi[\tau/\sigma])} Tr_=$$

$$\frac{ia : \phi}{i : (\psi \equiv \theta) \supset Bel_a^T(\phi[\psi/\theta])} Tr_{\equiv}$$

where $Tr_= [Tr_{\equiv}]$ is applicable only if both ϕ and $\tau = \sigma$ [$\psi \equiv \theta$] are closed formulae, and no formula the premiss $ia : \phi$ depends upon has index greater than i .

Figure 4 depicts the MultiContext architecture of MV^T ; notice that we added arrows corresponding to the Tr bridge rules.

Starting from Ω_ϵ , both the proofs of figure 3 and of figure 5 can be performed in MV^T . The last formula of the second proof corresponds to one of the possible transparent readings of the utterance $Bel_B("Bel_A("Corr(\iota P)")")$ that ϵ can perform.

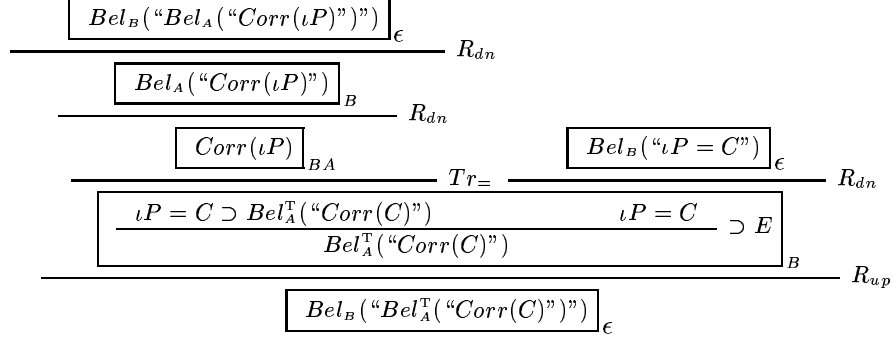


Fig. 5. Reasoning in MV^T

5.2 Modeling Transparency with LMS

Definition 4. An opaque model \mathbf{C} for MV^T is a transparent model if for every $\mathbf{c} \in \mathbf{C}$ and for every $i \in I$ and $a \in I_0$:

$$\begin{array}{l}
\text{if for all } \mathbf{c}' \in \mathbf{C}, \mathbf{c}'_i \subseteq \mathbf{c}_i \text{ implies } \mathbf{c}'_{ia} \models \phi, \text{ then} \\
\mathbf{c}_i \models \tau = \sigma \text{ implies } \mathbf{c}_i \models Bel_a^T(" \phi[\tau/\sigma] ") \quad (11)
\end{array}$$

$$\begin{array}{l}
\text{if for all } \mathbf{c}' \in \mathbf{C}, \mathbf{c}'_i \subseteq \mathbf{c}_i \text{ implies } \mathbf{c}'_{ia} \models \phi, \text{ then} \\
\mathbf{c}_i \models \psi \equiv \theta \text{ implies } \mathbf{c}_i \models Bel_a^T(" \phi[\psi/\theta] ") \quad (12)
\end{array}$$

(11) means that if ϕ is believed by ia , i.e. belongs to the mental image that i has of a , then i is able to perform a transparent reading of the formula $Bel_a(" \phi ")$ on the basis of the equality $\tau = \sigma$. (12) has an analogous meaning using equivalent formulae instead of terms.

We show how (11) allows the program to interpret Mr. B 's report of Mr. A 's belief $Corr(\iota P)$ using its own belief $\iota P = C$, instead of using A 's belief $\iota P = M$. Let us consider a transparent model \mathbf{C} for MV^T . Since any transparent model is also an opaque model, we obtain (7) from (5) as in the previous section. Constraint (11) between contexts \mathbf{B} and \mathbf{BA} , where ϕ is $Corr(\iota P)$, tells us that every compatibility sequence $\langle \mathbf{c}_\epsilon, \mathbf{c}_A, \mathbf{c}_B, \mathbf{c}_{AA}, \mathbf{c}_{AB}, \mathbf{c}_{BA}, \mathbf{c}_{BB} \dots \rangle$ is such that \mathbf{c}_B must satisfy

$$\boxed{B \quad (\iota P = C) \supset Bel_A^T("Corr(C)")} \quad (13)$$

From the fact that every \mathbf{c}_B satisfies $\iota P = C$ and from soundness of modus ponens in local models semantics [7], we obtain

$$\boxed{B \quad Bel_A^T(\text{"Corr}(C)\text{"})} \quad (14)$$

Finally, using constraint (4) as from (8) to (10), we obtain

$$\boxed{\epsilon \quad Bel_B(\text{"Bel}_A^T(\text{"Corr}(C)\text{"})\text{"})} \quad (15)$$

which corresponds to the transparent reading of Mr. *B*'s report of Mr. *A* belief.

Steps (13)–(15) are the model-theoretic counterpart of the salient steps of the proof in figure 5, where applications of constraint (11) correspond to applications of bridge rule $Tr_{=}$. The proof that MV^T systems are sound and complete with respect to the class of transparent models can be constructed following the methodology showed in [7].

6 Conclusions

The simple case study we have formalized in this paper is clearly an instance of a much more general phenomenon, potentially concerning any belief report where some reference is made to some individual. We argued that a proper treatment of this phenomenon does not involve a formalization of the *de re/de dicto* distinction, but rather a formalization of two possible attitudes on the reporter side towards another agent's beliefs. In the opaque reading, the reporter is assumed to just *quote* another agent's beliefs, whereas in the transparent reading the reporter is assumed to *translate* another agent's beliefs in his/her own words. As a consequence, the "same" definite description in the first case is meant to be used in the sense of the agent whose belief is reported, whereas in the second case is meant to be used in the sense of the reporter. This phenomenon is given a general (and quite natural) formalization by exploiting the multi-language features of MC systems.

The MC approach to belief contexts presupposes a deep change in the attitude toward the problem of formalizing beliefs (and reasoning about beliefs). Indeed, it requires to take seriously into account the fact that logical languages can be used to *ascribe beliefs to agents*, and not just to describe their beliefs from the point of view of an external (and – in general – omniscient) observer. The properties of locality and compatibility are motivated by this change of perspective, since they allow us not to hardwire in the logic assumptions which are implausible for an adequate logic of beliefs. For instance, we do not assume that the "same" linguistic token means the same thing for different agents, or that an agent can always understand what another agent meant to communicate. Misunderstandings happen all often in every day conversation, and an adequate formalization must explain how they are possible. We believe that our notion of views as contexts are a contribution toward such an explanation, and that it can throw a new light on some of the most difficult puzzles in formalizing beliefs (e.g. omniscience, failure of substitutivity, limitations of reasoning capabilities).

References

1. A. Bonomi. *Eventi mentali*. Il Saggiatore, 1983.
2. A. Cimatti and L. Serafini. Multi-Agent Reasoning with Belief Contexts: the Approach and a Case Study. In M. Wooldridge and N. R. Jennings, editors, *Intelligent Agents: Proceedings of 1994 Workshop on Agent Theories, Architectures, and Languages*, number 890 in Lecture Notes in Computer Science, pages 71–85. Springer Verlag, 1995. Also IRST-Technical Report 9312-01, IRST, Trento, Italy.
3. G. Frege. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik*, 100:25–50, 1892. English translation in [12].
4. C. Ghidini and L. Serafini. Distributed First Order Logics. Technical Report 9804-02, IRST, May 1998. Accepted for presentation at the First International Workshop on Labelled Deduction (LD'98). Freiburg, Germany. September 7 - 9, 1998.
5. E. Giunchiglia and F. Giunchiglia. Ideal and Real Belief about Belief. In *Practical Reasoning, International Conference on Formal and Applied Practical Reasoning, FAPR'96*, number 1085 in Lecture Notes in Artificial Intelligence, pages 261–275. Springer Verlag, 1996.
6. F. Giunchiglia. Contextual reasoning. *Epistemologia, special issue on I Linguaggi e le Macchine*, XVI:345–364, 1993. Short version in Proceedings IJCAI'93 Workshop on Using Knowledge in its Context, Chambéry, France, 1993, pp. 39–49. Also IRST-Technical Report 9211-20, IRST, Trento, Italy.
7. F. Giunchiglia and C. Ghidini. A Local Models Semantics for Propositional Attitudes. In *Proceedings of the 1st International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT-97)*, pages 363–372, Rio de Janeiro, Brazil, 1997. Also IRST-Technical Report 9607-12, IRST, Trento, Italy.
8. F. Giunchiglia and C. Ghidini. Local Models Semantics, or Contextual Reasoning = Locality + Compatibility. In *Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning (KR'98)*, pages 282–289, Trento, 1998. Morgan Kaufmann. Short version presented at the AAAI Fall 1997 symposium on context in KR and NL. Also IRST-Technical Report 9701-07, IRST, Trento, Italy.
9. F. Giunchiglia and L. Serafini. Multilanguage hierarchical logics (or: how we can do without modal logics). *Artificial Intelligence*, 65:29–70, 1994. Also IRST-Technical Report 9110-07, IRST, Trento, Italy.
10. F. Giunchiglia, L. Serafini, E. Giunchiglia, and M. Frixione. Non-Omniscient Belief as Context-Based Reasoning. In *Proc. of the 13th International Joint Conference on Artificial Intelligence*, pages 548–554, Chambéry, France, 1993. Also IRST-Technical Report 9206-03, IRST, Trento, Italy.
11. J. Hintikka. *Knowledge and Belief*. Cornell University Press, Ithaca, NY, 1962.
12. A. P. Martinich. *The philosophy of language*. Oxford University Press, 1985.
13. D. Prawitz. *Natural Deduction - A proof theoretical study*. Almqvist and Wiksell, Stockholm, 1965.
14. B. Russell. On denoting. *Mind*, XIV, 1905.