We can thus restrict ourselves to normal form deductions. The proof is again by induction on the length of the deduction. All the deductions not involving R_{down} and R_{up} are trivially translated and not considered. As we are considering normal form deductions, we cannot have branches with applications of R_{up} before the application of R_{down}. Moreover, since all the deductions start from wffs of index ϵ and end with a wff with index ϵ, there must be a branch where the application of R_{down} must be followed by the application of R_{up}. Thus any normal deduction of $\langle A^*, \epsilon\rangle$ from $\langle \Gamma^*, \epsilon\rangle$ involving the application of R_{down} and R_{up} must have the following shape:

$$\dfrac{\begin{array}{c}\Sigma_1\\ \langle B_i(``C_1^*"), \epsilon\rangle\end{array}}{\langle C_1^*, i\rangle} \quad \cdots \quad \dfrac{\begin{array}{c}\Sigma_n\\ \langle B_i(``C_n^*"), \epsilon\rangle\end{array}}{\langle C_n^*, i\rangle}$$

$$\begin{array}{c}\Sigma\\ \dfrac{\langle A^*, i\rangle}{\langle B_i(``A^*"), \epsilon\rangle}\\ \Sigma_{n+1}\end{array}$$

where Σ is a deduction using only $\rho(i)$. The above $MB(L,\rho)$ deduction corresponds to the $B(L,\rho)$ proof of the sequent $[S_i]\Gamma \Rightarrow [S_i]A$ from $\Gamma \vdash_{\rho(i)} A$, which is exactly the attachment rule and thus a valid inference rule of $B(L,\rho)$. $\triangle$

Appendix A: the proof of theorem 1

While reading the proof, it must be taken into account that Konolige uses a sequent version of the semantic tableaux method while we use ND. This is irrevelant from the point of view of the correctness of the proof, but it must be kept in mind not to generate confusion. The difference between the semantic tableaux method and ND does not play any role here. Instead, as far as the relation between a sequent calculus and a ND calculus is concerned, the former is given the following natural interpretation in terms of the second. When we come to application of a sequent calculus succedent rule, we enlarge the corresponding ND proof at the bottom applying the corresponding I-rule. When we come to the application of an antecendent rule, we enlarge the corresponding ND proof at the top, applying the corresponding E-rule.

Proof[Hinted]
($\Rightarrow$) The proof is by inductively transforming a proof for the sequent $\Gamma \Rightarrow A$, in a deduction of $\langle A^*, \epsilon\rangle$, from $\langle \Gamma^*, \epsilon\rangle$. The induction steps are obvious and result in a translation of sequent calculus proofs into ND deductions (following what discussed above, before the proof). Complications rise in the base case. For how $B(L,\rho)$ has been defined, there are two possible bottom sequents [10] and thus two base cases. One possible bottom sequent is "$\Gamma, A \Rightarrow \Delta, A$". Its obvious translation is a deduction of $\langle A^*, \epsilon\rangle$ from $\langle A^*, \epsilon\rangle$. The second bottom formula is the premise of the attachment rule "$\Lambda,\ [S_i]\Gamma \Rightarrow [S_i]A,\ \Delta$" (where $\Gamma = \{C_1, ..., C_n\}$). The proof ends only if $\Gamma \vdash_{\rho(i)} A$ holds (in other words, with an application of the attachment rule). In this case both the application of the attachment rule and the deduction in the i-th deduction structure must be translated into a $MB(L,\rho)$ deduction. Their translation is as follows:

$$\rho'(i)\ \frac{R_{down}\dfrac{\langle B_i(\text{“}C_1^*\text{”}),\ \epsilon\rangle}{\langle C_1^*,\ i\rangle}\ \ldots R_{down}\dfrac{\langle B_i(\text{“}C_n^*\text{”}),\ \epsilon\rangle}{\langle C_n^*,\ i\rangle}}{R_{up}\dfrac{\langle A^*,\ i\rangle}{\langle B_i(\text{“}A^*\text{”}),\ \epsilon\rangle}}$$

where the line labeled with $\rho'(i)$ represents the deduction performed (using $\rho(i)$) to derive $\langle A^*, \epsilon\rangle$ from $\langle C_1^*, \epsilon\rangle ..., \langle C_n^*, \epsilon\rangle$. $\langle \Gamma^*, i\rangle \vdash_{\rho'(i)} \langle A^*, i\rangle$ is the $MB(L,\rho)$'s version of "$\Gamma \vdash_{\rho(i)} A$". $\langle \Gamma^*, i\rangle \vdash_{\rho'(i)} \langle A^*, i\rangle$ is a valid $MB(L,\rho)$ deduction by hypothesis.
($\Leftarrow$) The basic idea underlying this part of the proof is that $MB(L,\rho)$ deductions can be proved to have a normal form [11]. Our notion of normal form is a generalization of Prawitz' one in the sense that R_{up} (R_{down}) is seen as B_i introduction (elimination). One property of $MB(L,\rho)$ normal deductions is that any of their branches (our notion of branch is the same as in [Pra65]) can be separated in three (possibly empty) subparts : the first involving a sequence of eliminations in L_ϵ, the second involving deduction using only $\rho(i)$ (for a given i) the third involving a sequence of introductions in L_ϵ. The first and second step are possibly separated by an application of R_{down}, the second and the third, possibly, by an application of R_{up}.

[10]Remember footnote 3.

[11]A deduction is in normal form when it does not contain any maximum formulas, *ie.* no formula that is the major premiss of an elimination rule and the consequence of an introduction rule.

[Kon82] K. Konolige. A First-Order Formalisation of Knowledge and Action for a Multi-Agent Planning System. In *Machine Intelligence 10*, 1982.

[Kon84] K. Konolige. *A deduction model of belief and its logics.* PhD thesis, Stanford University CA, 1984.

[Pra65] D. Prawitz. *Natural Deduction - A proof theoretical study.* Almquist and Wiksell, Stockholm, 1965.

[RL86] J. Des Rivieres and H.J. Levesque. The consistency of syntactical treatments of modality. In J.Y. Halpern, editor, *Proceedings of the 1986 Conference on theoretical Aspects of Reasoning about knowledge*, pages 115–130, 1986.

[WB79] Y. Wilks and J. Biem. Speech acts and multiple environments. In *Proc. of the 6th International Joint Conference on Artificial Intelligence*, 1979.

[Wey80] R.W. Weyhrauch. Prolegomena to a Theory of Mechanized Formal Reasoning. *Artif. Intell.*, 13(1):133–176, 1980.

6 Conclusion and final remarks

We have presented the general definition of ML-system. This notion, together with the notion of bridge rules, is new and allows us to treat elegantly and naturally situations which before were treated by using modal operators. We have concentrated on propositional attitudes and proved various equivalence and consistency results., We have also claimed that ML-systems are easily implementable on a computer and very briefly introduced a system (`GETFOL`) which gives the user the ability to define arbitrary ML-systems.

At a closer look, an ML-system seems to have a number of advantages over modal and Konolige's belief logics. First, we do not need to extend the syntax to consider modal operators whose epistemological meaning is not obvious (at least to the authors). Second, as we have many distinct languages, we can really choose what can be said and where. Thus, for instance, the case where the external and the internal languages are distinct is dealt with very naturally and without many complications.

References

[Fau85] G. Fauconnier. *Mental Spaces: aspects of meaning construction in natural language*. MIT Press, 1985.

[Fef62] S. Feferman. Transfinite Recursive Progressions of Axiomatic Theories. *Journal of Symbolic Logic*, 27:259–316, 1962.

[GS89] F. Giunchiglia and A. Smaill. Reflection in constructive and non-constructive automated reasoning. In H. Abramson and M. H. Rogers, editors, *Proc. Workshop on Meta-Programming in Logic Programming*, pages 123–145. MIT Press, 1989. Also IRST-Technical Report 8902-04 and DAI Research Paper 375, University of Edinburgh.

[GW88] F. Giunchiglia and R.W. Weyhrauch. A multi-context monotonic axiomatization of inessential non-monotonicity. In D. Nardi and P. Maes, editors, *Meta-level architectures and Reflection*, pages 271–285. North Holland, 1988. Also MRG-DIST Technical Report 9105-02, DIST, University of Genova, Italy.

[Hal86] J.Y. Halpern. Reasoning about knowledge, an overview. In *Proceedings of the 1986 Conference on theoretical aspects of reasoning about knowledge*, pages 1–17. Morgan Kaufmann Publ., Inc, 1986.

[HM85] 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.

[JL83] J. N. Johnson-Laird. *Mental Models*. Harvard University Press, 1983.

can be given, for instance, by using a metatheoretic predicate expressing the consistency of a wff with a set of hypotheses (along the lines described in [GW88]). On the other hand, it is well known that the set of wffs consistent with a given theory is in general not recursively enumerable. To avoid the problem let us restrict to the case of theories with a decidable deducibility relation. In these cases the possibly infinite set $\{\langle \neg D_i(\text{“}\Gamma\text{”}, \text{“}A\text{”}),\ i\rangle : \langle \Gamma,\ i\rangle \not\vdash_{\tau(i)} \langle A,\ i\rangle, i \in I\}$ is recursive and can be given as a set of axioms (possibly by using axiom schemas).

A second extension considers **common beliefs**. Common beliefs are believed by everyone and everyone believes that they are common beliefs (see [HM85] for some related work on this topic). They can be axiomatized by introducing a new predicate symbol $CB(.)$ and by adding $\langle CB(\text{“}A\text{”}) \rightarrow B_i(\text{“}A\text{”}),\ \epsilon\rangle$ and $\langle CB(\text{“}A\text{”}) \rightarrow B_i(\text{“}CB(\text{‘}A\text{’})\text{”}),\ \epsilon\rangle$ as axioms of the external language or by adding the following two rules

$$CB_1 \frac{\langle CB(\text{“}A\text{”}),\ \epsilon\rangle}{\langle A,\ i\rangle} \qquad CB_2 \frac{\langle CB(\text{“}A\text{”}),\ \epsilon\rangle}{\langle CB(\text{“}A\text{”}),\ i\rangle} \tag{3}$$

to the deduction machinery. The two approaches are technically equivalent.

Let us now define the (twice) saturated ML-system $MBK_s^{+s}(L,\tau)$ with language $L(\{B_i\}_{i\in I} \cup \{D_i\}_{i\in\omega} \cup \{CB\})$ obtained by extending $MBK_s(L,\tau)$ to deal with circumscriptive ignorance and common beliefs as follows [8] [9]:

Definition 5 ($MBK_s^{+s}(L,\tau)$) *: $MBK_s^{+s}(L,\tau)$ is a ML-system obtained from $MBK_s(L,\tau)$ by modifying the set of axioms and of inference rules as follows:*

1. *for each view $\nu \in I^*$ $A_\nu = \{\langle \neg D_i(\text{“}\Gamma\text{”}, \text{“}A\text{”}),\ \nu\rangle : \langle \Gamma,\ \nu i\rangle \not\vdash_{\tau(\nu i)} \langle A,\ \nu i\rangle\}$;*

2. *each view has a set of ND rules for classical propositional logic. The bridge rules are*

$$R_{down.\nu i} \frac{\langle B_i(\text{“}A\text{”}),\ \nu\rangle}{\langle A,\ \nu i\rangle} \quad R_{up.\nu} \frac{\langle A,\ \nu i\rangle}{\langle B_i(\text{“}A\text{”}),\ \nu\rangle}$$

$$Der\ \frac{\begin{matrix}[\langle A_1,\ \nu i\rangle]\ \cdots [\langle A_n,\ \nu i\rangle] \\ \langle B,\ \nu i\rangle\end{matrix}}{\langle D_i(\text{“}A_1\text{”},\ \cdots,\ \text{“}A_n\text{”}, \text{“}B\text{”}),\ \nu\rangle}$$

for every view $\nu \in I^*$ with the usual restrictions on the indexes. and

$$CB_1 \frac{\langle CB(\text{“}A\text{”}),\ \nu\rangle}{\langle A,\ \nu i\rangle} \quad CB_2 \frac{\langle CB(\text{“}A\text{”}),\ \nu\rangle}{\langle CB(\text{“}A\text{”}),\ \nu i\rangle}$$

MBK_s^{+s} can be proved to be consistent and equivalent to Konolige's BK_s^+ (with the restriction that the agents' deducibility relations must be recursive). MBK_s^{+s} is expressive enough to represent the solution of the not-so-wise-man problem [Kon84]. An implementation (not described here for lack of space) of the solution of such problem in the `GETFOL` mechanization of MBK_s^{+s} has been performed.

[8]In $MBK_s(L,\tau)$ the theory in the external language is identical to that to any other view. To simplify the notation we consider ν to belong to the set I^* of the (possibly empty) finite sequences of I and assign the external language the empty sequence.

[9]The second saturation is due to the fact that all the views have been added circumscriptive ignorance and common beliefs.

Let us now go back to Konolige's logics. By applying a process analogous to that used to obtain MBK from MB, we can obtain the system $BK(L,\tau)$ from $B(L,\tau)$ (see [Kon84] for a detailed description of how this can be done). In particular consider definition 2 and substitute items 1 and 2 with the following:

1. The attachment rule must be defined for any view $\nu \in \overline{I}$ (see definition 4). In other words the attachment rule becomes

$$A_K \ \frac{\Lambda,\ [S_i]\Gamma \Rightarrow [S_i]A,\ \Delta}{\Gamma \Rightarrow_i A}$$

2. The belief derivation operator is, for each view ν, the sequent calculus translation of the set of ND L_ν-rules $\tau(\nu)$ (notice that the sequent of shape "$\Gamma,\ A \Rightarrow A,\ \Delta$" is derivable and the *Cut-rule* is a derived inference rule).

Moreover, analogously to $MBK_s(L,\tau)$, let $BK_s(L,\tau)$ be the saturated version of $BK(L,\tau)$. The following results hold.

Theorem 2 *: $\vdash_S \Gamma \Rightarrow A$ if and only if $\langle \Gamma^*,\ \epsilon\rangle \vdash_{MS} \langle A^*,\ \epsilon\rangle$, where S, MS can be respectively $B(L,\rho)$, $MB(L,\rho)$ or $BK(L,\tau)$, $MBK(L,\tau)$ or $BK_s(L,\tau)$, $MBK_s(L,\tau)$.*

Corollary 3 *: $MBK(L,\tau)$ and $MBK_s(L,\tau)$ are consistent.*

5 An Extension

Let us first consider **circumscriptive ignorance**: an agent *does* believe what it is able to derive and *does not* believe what it is not able to derive. Notice that an agent given circumscriptive ignorance is able to derive something from ignorance, *ie.* the fact that something cannot be derived is the basis for not believing it (for stating the negation of its belief). An ML-system can be extended to capture circumscriptive ignorance as follows. For any natural n we introduce a set of $n+1$-ary predicate symbols D_i^n (from now on we drop the superscript n), where the meaning of $D_i(\text{"}A_1\text{"},\ \cdots,\ \text{"}A_n\text{"},\ \text{"}A\text{"})$ is that it is possible to derive $\langle A,\ i\rangle$ from $\langle A,\ 1\rangle, ..., \langle A,\ n\rangle$ by $\rho(i)$ [7]. For any agent i ($i \in I$), the ability to believe that a wff can be derived from a set of wffs (when this is the case) can be captured with the following rule:

$$Der \ \frac{\begin{array}{c}[\langle A_1,\ i\rangle]\ \cdots [\langle A_n,\ i\rangle]\\ \langle B,\ i\rangle\end{array}}{\langle D_i(\text{"}A_1\text{"},\ \cdots,\ \text{"}A_n\text{"}, \text{"}B\text{"}),\ \epsilon\rangle}$$

with the restriction that any not discharged assumption has index ϵ or is one of the $\langle A_1,\ i\rangle,\ \cdots,\ \langle A_n,\ i\rangle$ (note that $\langle A_1,\ i\rangle,\ \cdots,\ \langle A_n,\ i\rangle$ are discharged by Der).

The formalization of the ability of deriving the negation of one belief from the inability of deriving something is far more problematic. A possible multicontext version of

[7] We write $D_i(\text{"}\Gamma\text{"}, \text{"}A\text{"})$ to mean $D_i^n(\text{"}A_1\text{"},\ \cdots \text{"}A_n\text{"},\ \text{"}A\text{"})$, where $\Gamma = \{A_1,\ \ldots,\ A_n\}$.

4 Some particular cases

Various versions of $MB(L,\rho)$ can be built simply by choosing a different set of $\rho(i)$'s. Both the external and the internal language allow to express nested beliefs (*ie.* beliefs about beliefs), for instance B_{John}("B_{Sue}("A")") means that *John* believes that *Sue* believes that A. Thus an obvious choice is that the agents' theories are themselves one of the logics of the logic family $MB(L,\rho)$. Intuitively the idea is that the theory of each agent propagates down recursively (via the use of bridge rules) into new theories. Any theory can be univocally identified by a (possible empty) sequence of indexes $\nu = i_1 ... i_n$, called *view*. The theory associated with any view $\nu = i_1 i_2 ... i_n$ represents the i_1 beliefs about i_2 beliefs about ... about i_n beliefs. A further restriction is that each view uses ND rules. These considerations suggest the definition of the following ML-sistem (if $\nu = i_1 \ldots i_n$ is a (not empty) sequence of elements of a set I and $i \in I$ "νi" denotes the sequence $i_1 \ldots i_n i$):

Definition 4 ($MBK(L,\tau)$) *: Let I be a countable set of indexes, I^+ the set of the not empty finite sequences of elements of I and $\overline{I} \subseteq I^+$. For any $\nu \in I^+$, let $\tau(\nu)$ be a set of ND rules for the language $L(\{B_i\}_{i \in I})$ admitting a normal form. Then $MBK(L,\tau)$, $MBK(L,\tau) = \langle \{L_\nu\}_{\nu \in I^+ \cup \{\epsilon\}},\ \{A_\nu\}_{\nu \in I^+ \cup \{\epsilon\}},\ \tau' \rangle$ ($\epsilon \notin I$), is an ML-system such that*

1. *For any $\nu \in I^+ \cup \{\epsilon\}$, L_ν is $L(\{B_i\}_{i \in I})$, $A_\mu = \emptyset$;*
2. *τ', is the set of ND L_ϵ-rules (for classical propositional logic) union the set of ND L_ν-rules $\tau(\nu)$ ($\nu \in I^+$) union the bridge rules:*

$$R_{down.i}\frac{\langle B_i(\text{“}A\text{”}),\ \epsilon\rangle}{\langle A,\ i\rangle} \qquad R_{up.i}\frac{\langle A,\ i\rangle}{\langle B_i(\text{“}A\text{”}),\ \epsilon\rangle}$$

for every $i \in I$ and

$$R_{down.\nu i}\frac{\langle B_i(\text{“}A\text{”}),\ \nu\rangle}{\langle A,\ \nu i\rangle} \qquad R_{up.\nu i}\frac{\langle A,\ \nu i\rangle}{\langle B_i(\text{“}A\text{”}),\ \nu\rangle}$$

for every view $\nu \in \overline{I}$ with the usual restriction on the application of the R_{up}-rules.

Notice that it is not required that $\tau(\nu)$ be complete for the propositional calculus nor that $R_{down.\nu i}$,$R_{up.\nu i}$ are defined for any pair of views $\nu, \nu i$. $MBK_s(L,\tau)$ ($MBK(L,\tau)$ saturated) is obtained from $MBK(L,\tau)$ by asking the two requirements above.

$MBK(L,\tau)$ and $MBK_s(L,\tau)$ can be proved to be equivalent to various systems proposed in the modal logic and belief literature. In particular, if I is a singleton then $MBK_s(L,\tau)$ is equivalent to the modal system K. Moreover if K_I is the multimodal K system with language $L(\{M_i\}_{i \in I})$ where all the modal operators M_i satisfy the modal system K properties (in other words, necessitation and the axiom "$M_i A \rightarrow B \wedge M_i A \rightarrow M_i B$"), then the following fact holds.

Corollary 2 $\Gamma \vdash_{K_I} A \quad \Longleftrightarrow \quad \langle \Gamma^*,\ \epsilon\rangle \vdash_{MBK_s(L,\tau)} \langle A^*,\ \epsilon\rangle$

(1) If A is a propositional constant then $A^* = A$;

(2) $(.)^*$ is distributive over the logical connectives and

(3) $([S_i]A)^* = B_i(\text{“}A^*\text{”})$.

$(.)^*$ is an isomorphism and it has an inverse. Notice that $(.)^*$ preserves the intuitive meaning of the wffs. Indeed, for instance, both $[S_i]A$ and $B_i(\text{“}A^*\text{”})$ mean that the agent i believes A. Notationally, if Γ is a set of $L(\{[S_i]\}_{i\in I})$-wffs, $\langle\Gamma^*, \epsilon\rangle =_{def} \{\langle B^*, \epsilon\rangle : B \in \Gamma\}$.

Theorem 1 *:* $\vdash_{B(L,\rho)} \Gamma \Rightarrow A \iff \langle\Gamma^*, \epsilon\rangle \vdash_{MB(l,\rho)} \langle A^*, \epsilon\rangle$.

The proof of theorem 1 is in Appendix A.

Corollary 1 *: $MB(L, \rho)$ is consistent.*

Proof[Hinted]: If $MB(L, \rho)$ were inconsistent, because of theorem 1, so would be $B(L, \rho)$. But this is false as $B(L, \rho)$ has a model [Kon84]. $\triangle$

Some observations.

Even if theorem 1 tells us that, in a sense, we could indifferently use ours or Konolige's logic, our claim is that our approch leads to a more natural representation of belief. The intuition, that the sets of beliefs associated with the agents and the external observer are separate and distinct from one another, is explicitly represented in the formalism. On the technical side this allows us to treat the "propagation" of results from one language to another purely syntactically: we do not need any form of attachment rule between the syntax and the semantics.

Before developing the work described in [Kon84] and analyzed above, Konolige proposed a syntactic approach to belief [Kon82] where, as in our case, belief was represented as a first order metatheoretic predicate. He later gave up this approach and in his thesis [Kon84] motivated this by writing that the first order approach leads to too much complexity. The simplicity and naturalness of the definition of $MB(L, \rho)$ and of theorem 1 shows that, if we give up the hypothesis of having only one language for representing the agents' beliefs (as we have done), the treatment becomes very natural and arguably simpler than the modal one. Notice that here, contrarily to what happens in [Kon82], we restrict ourselves only to the propositional case. The generalisation to first order is under development but it does not seem to lead to a high level of complexity in the notation, theory or implementation.

The equivalence result stated by theorem 1 is very different from the equivalence result stated in [RL86]. In this work, the authors stick to the hypothesis of having an unique (first order) language and obtain the equivalence by restricting the set of well formed formulas considered.

To make precise this intuitive introduction to the ML-system corresponding to the picture above, the first step is the definition of the external and the agents' languages. To simplify things, following Konolige, we suppose that all the languages have the same syntactic structure (in other words, they define the same set of wffs). If $\{M_i^{n_i}\}_{i\in I}$ is the set of predicates corresponding to the modal operators (where n_i is M_i's arity, dropped from now on), then $L(\{M_i\}_{i\in I})$ is the first order metatheoretic language obtained from a propositional language L and $\{M_i\}_{i\in I}$ as follows [5]:

- $L_0 = L \cup \{M_i\}_{i\in I}$;
- $L_{n+1} = L_n$ supplied with the names of the L_n-wffs (where, if A is a L-wff, then "A" is its name), that are not L_m-wffs for some $m < n$;
- $L(\{M_i\}_{i\in I}) = \bigcup_{n\in\omega} L_n$.

Note that different $L(\{M_i\}_{i\in I})$-wffs have different names and that $L(\{M_i\}_{i\in I})$ contains a name for each $L(\{M_i\}_{i\in I})$-wff.

We can thus define $MB(L,\rho)$, a system provably equivalent to $B(L,\rho)$ [6].

Definition 3 ($MB(L,\rho)$) *: Let I be a countable set of indexes and, for every $i \in I$, $\rho(i)$ a set of inference rules for the language $L(\{B_i\}_{i\in I})$. Then $MB(K,\rho)$, $MB(L,\rho) = \langle \{L_i\}_{i\in I\cup\{\epsilon\}}, \{A_i\}_{i\in I\cup\{\epsilon\}}, \rho' \rangle$ ($\epsilon \notin I$), is a ML-system such that :*

1. *for every $i \in I \cup \{\epsilon\}$, L_i is $L(\{B_i\}_{i\in I})$, $A_i = \emptyset$;*
2. *ρ' is composed of a set of (complete) classical ND L_ϵ-rules, the agents' rules $\rho(i)$ ($i \in I$) and the two bridge rules:*

$$R_{up}\frac{\langle A,\ i\rangle}{\langle B_i(\text{“}A\text{”}),\ \epsilon\rangle} \qquad R_{down}\frac{\langle B_i(\text{“}A\text{”}),\ \epsilon\rangle}{\langle A,\ i\rangle}$$

where $R_{up.}$ is applicable only if the index of any not discharged assumption is ϵ.

I's cardinality is the number of the agents. ϵ is the index associated with the external observer. B_i is the translation of $[S_i]$. The single theories are arranged in a tree of depth 1. The root is the theory of the external observer, the i-th leaf represents the beliefs of the agent i, each arc is labeled by two bridge rules (see picture above).

The definition of the equivalence between the two systems consists of two main steps: (i) definition of the mapping between the languages of the systems and (ii) proof that derivability is invariant under this mapping.

Let us define a mapping $(.)^*$ from $L(\{[S_i]\}_{i\in I})$-wffs to $L(\{B_i\}_{i\in I})$-wffs as follows:

[5]Notice that we (deliberately) use the same notation for modal and first order languages. The context always makes clear the version considered.

[6]Note that the "ρ" in $MB(L,\rho)$ is different from that in $B(L,\rho)$. The considerations done in footnote 5 apply here too.

Each agent is indentified by a *deduction structure* $\langle B, \mathcal{R}\rangle$ where B is the set of *base beliefs* (a set of sentences in an *internal* (logical) *language* L), and $\mathcal{R}$ is a set of inference rules. The *belief set* of an agent is the transitive closure of $\mathcal{R}$ over B, *ie.* $\{A \in L : B \vdash_{\mathcal{R}} A\}$. Thus the interpretation of Konolige's modal logics is composed of the interpretation of the propositional language L, supplied with a set of deduction structures $\{d_i = \langle B, \rho(i)\rangle : B \subseteq L^B,\ i \in I\}$. For instance $[S_i]A$ is true if A belongs to the belief set of d_i.

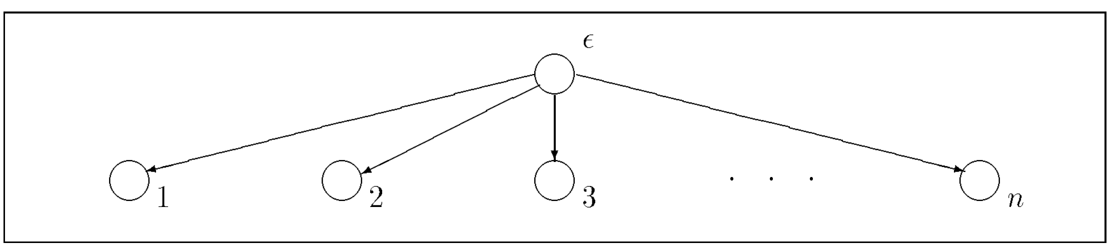
Notationally, if I is a countable set of indexes, by $L(\{M_i\}_{i\in I})$ we mean the modal propositional language obtained by adding the set of modal operators $\{M_i\}_{i\in I}$ to the propositional language L. Konolige's basic system is the system $B(L, \rho)$ whose language L^B is $L(\{[S_i]\}_{i\in I})$, with I countable. ρ is a function that assigns to every agent $i \in I$ a set of deduction rules $\rho(i)$ with correspondent deducibility relation $\vdash_{\rho(i)}$ [2]. $B(L, \rho)$ has no axioms, its deductive machinery is as follows:

Definition 2 (Konolige [Kon84]) *: $B(L, \rho)$'s deductive machinery has three components:*

1. *The propositional complete rules T_0 of the block tableaux;*
2. *the attachment rule* [3] A *:* $\dfrac{\Sigma, [S_i]\Gamma \ \Rightarrow\ [S_i]A, \Delta}{\Gamma \vdash_{\rho(i)} A}$;
3. *a belief derivation operator $\vdash_{\rho(i)}$ for each agent S_i.*

The sequent "$[S_i]\Gamma \Rightarrow [S_i]A$" holds **if and only if** $\Gamma \vdash_{\rho(i)} A$. Notice that the attachment rule "links" derivations in the agents's deduction structures to the provability of (certain) sequents in the logic [4]. The attachment rule allows the **"propagation" of results from a syntactic model into its logic**.

The ML-systems provably equivalent to Konolige's logics of belief have three main components. The first is the theory ϵ of the external observer (written in the external language) with a first order predicate for each modal operator. The second is a set of theories $1, 2, ..., n$ (written in the internal language) which correspond to the agents' deduction structures. Finally, the third is a set of bridge rules, each axiomatizing **syntactically** the steps from ϵ to each of the theories i. The following picture formalizes what we have just said.



[2] $B(L, \rho)$ is based on a sequent version of Hintikka and Kripke's semantic tableaux method. "$\Rightarrow$" is the symbol denoting the sequent. $[S_i]\Gamma =_{def} \{[S_i]\gamma : \gamma \in \Gamma\}$; thus, for instance, the intended meaning of "$[S_i]\Gamma \Rightarrow [S_i]A$" is that *if S_i believes every γ in Γ then he believes A.*

[3] Konolige's rules are upside down with the conclusion above and the premises below the horizontal line.

[4] This operation is a variation of the idea of attachment as it is implemented in the `FOL` system [Wey80].

Two examples of rules which can be seen as bridge rules, discussed in detail in [GS89], are **reflection down** (a form of reflection) and **reflection up** (a form of necessitation):

$$R_{down}\ \frac{\vdash_{T_1}\ Prov_{T_2}(\text{“}\phi\text{”})}{\vdash_{T_2}\ \phi} \qquad\qquad R_{up}\ \frac{\vdash_{T_2}\ \phi}{\vdash_{T_1}\ Prov_{T_2}(\text{“}\phi\text{”})}$$

These two inference rules are paradigmatic. Similar ones have been used in a lot of (the informally described) work in metatheoretic theorem proving (see [GS89]) even if in most cases (with the notable exception of [Wey80]) T_1 and T_2 shared the same language. Their multilanguage version, if considered as bridging deduction between the languages i and $i+1$ is:

$$R_{down,i}\ \frac{\vdash \langle Prov_i(\text{“}\phi\text{”}),\ i+1\rangle}{\vdash \langle \phi,\ i\rangle} \qquad\qquad R_{up,i+1}\ \frac{\vdash \langle \phi,\ i\rangle}{\vdash \langle Prov_i(\text{“}\phi\text{”}),\ i+1\rangle}$$

Very similar versions are used in the formal systems which are proved equivalent to Konolige's logics (see later) and also in the definition of ML-systems which are equivalent to the various modal logics (a paper on this last topic is forthcoming).

In the past it has been longly spoken, both in the artificial intelligence and in the mathematical logic community of hierarchies of metatheories, of self-reflective theories and so on ([Fef62] is, in our opinion, one of the most significant examples of this tradition). Our notion of formal system is novel for two reasons. First, multiple languages can be defined inside the formal system itself. In the past, the hierarchy was often seen as the incremental extension of the same theory (as it happens in [Fef62]). There was only one theory at the time with no notion of multiple theories with distinct languages. The metatheory was always a (conservative or not) extension of the object theory. A similar behaviour can be obtained with an ML-system with no bridge rules. Second, just because we have multiple languages inside the same formal system, we can formally define bridge rules as part of the deductive machinery of an ML-system. The advantage is that in this case the theory-to-theory interaction can be formalized **in** the ML-system itself and not in its (informal) metatheory. This becomes particularly relevant if we consider that a lot of the latest work in AI seems to push towards the use of system with (implicitly or explicitly) separated languages.

As a final observation, a system which implements these ideas and gives the user the ability of define arbitrary ML-systems (with arbitrary bridge rules) already exists and it is a re-implementation/ extension of the `FOL` system [Wey80], called `GETFOL`. For instance, the ML-system described in section 5 has been implemented in `GETFOL` and used to prove the not-so-wise-man problem [Kon84].

3 Konolige's versus multilanguage systems of belief

Konolige's logics represent situations where a set of entities (called agents) are defined which have a representation of the external world, of his and other agents' beliefs, and are able to reason about it. The language of these logics, also called the *external language*, is modal. Modalities are used to state facts about the agents' beliefs; thus, for instance, $[S_i]A$ means that the agent S_i believes A.

Definition 1 (Multilanguage Formal System) *: Let I be a set of indices, $\{L_i\}_{i\in I}$, a family of languages and $\{A_i\}_{i\in I}$ a family of sets of L_i-wffs* [1]*. A* **Multi-Language Formal System** *(***ML-System***) Σ is a triple $\langle\{L_i\}_{i\in I}, \{A_i\}_{i\in I}, \Delta\rangle$ where $\{L_i\}_{i\in I}$ is the* **Family of Languages***, $\{A_i\}_{i\in I}$ is the* **Family of Axioms** *and Δ is the* **Deductive machinery** *of Σ.*

We write $\langle A,\ i\rangle$ to mean A and that A is a L_i-wff. Note that, for some $i, j \in I$, we may have $L_i = L_j$. Note also that a natural deduction (ND from now on) system is a particular case of an ML-system (when I is a singleton).

The deduction machinery Δ, is a set of *inference rules*, a la Prawitz, with corresponding *deduction rules* which describe how inference rules may be used to produce deductions. Informally, we write inference rules as:

$$\iota : \ \frac{\langle A_0,\ i_0\rangle \ \ldots \ \langle A_{n-1},\ i_{n-1}\rangle}{\langle A_n,\ i_n\rangle} \tag{1}$$

or as:

$$\delta : \frac{\langle A_0,\ i_0\rangle \ \ldots \ \langle A_{n-1},\ i_{n-1}\rangle \ \ \begin{matrix}[\langle B_0,\ j_0\rangle] \\ \langle C_0,\ k_0\rangle\end{matrix} \ \ \begin{matrix}\ldots \\ \ldots\end{matrix} \ \ \begin{matrix}[\langle B_{m-1},\ j_{m-1}\rangle] \\ \langle C_{m-1},\ k_{m-1}\rangle\end{matrix}}{\langle A,\ l\rangle} \tag{2}$$

(2) represents a rule δ discharging the assumptions $\langle B_0,\ j_0\rangle, \ldots, \langle B_{m-1},\ j_{m-1}\rangle$.

Deductions are trees of wffs built starting from a finite number of assumptions and applying a finite number of inference rules. $\langle A,\ i\rangle$ is *derivable* from a set of wffs Γ in a ML-system Σ ($\Gamma \vdash_\Sigma \langle A,\ i\rangle$) if there is a deduction with bottom wff $\langle A,\ i\rangle$ whose undischarged assumptions are in Γ. $\langle A, i\rangle$ is a theorem in Σ ($\vdash_\Sigma \langle A, i\rangle$) if it is derivable from the empty set.

An in depth discussion of the basic ideas underlying the system is out of the goals of the paper. Some observations are, on the other hand, worthwhile.

Each language is associated with a theory. Intuitively the major advantage of having distinct theories is that it is possible to keep under control what can be said in one language. Thus, for instance, it is possible to state facts *about* a theory in a language different from that used for stating facts *in* the theory.

In an ML-system theories are not independent of one another, since we allow the definition of a new kind of inference rules, called **bridge rules**, whose premises and conclusions belong to distinct languages. Bridge rules control the propagation of results among theories. We call L_i-rules the rules which are not bridge rules, where L_i is the language of their premises and conclusions. The usual monolingual (ND) inference rules can be transformed in L_i-rules. Thus the multilanguage version of conjunction introduction:

$$\wedge I \ \ \frac{A \ \ B}{A \wedge B}$$

in the language i is:

$$\wedge I_i \ \ \frac{\langle A,\ i\rangle \ \ \langle B,\ i\rangle}{\langle A \wedge B,\ i\rangle}$$

[1]A **family** is a set with repetitions. $\{L_i\}_{i\in I}$ and $\{A_i\}_{i\in I}$ can be constructed as the codomain of two functions f_L and f_A respectively with domain I.

(as it is well known, from a formal point of view, frame-like languages, can be seen as syntactic sugar of (subclasses of) first order logic); one example in this area is [WB79].

Both approaches share the same underlying hypothesis that the representation language is unique. The aim of the research (partially) described in this paper is to propose a new formalism which makes effective use of multiple **distinct first order languages**. Each language is associated with it's own theory: this allows us to structure in a more effective way the knowledge base and the reasoning about it. Reasoning in a theory/ language is in general not independent of reasoning in the other theories as we allow the definition of a new kind of inference rules, called **bridge rules**, whose premises and conclusions belong to distinct languages.

A representation of knowledge as a set of multiple languages/theories leads to more efficent implementations (for instance, the search space is smaller) but it has also epistemological motivations. This proposal seems in fact very much in agreement with a lot of the work done in cognitive science. Thus, for instance, Wilks in [WB79] advocates the use of distinct sets of beliefs; Faucounnier [Fau85] has a mental space theory which uses environment like entities while Johnson-Laird [JL83], in an application of his theory to propositional attitudes, uses explicit, nested groups of representational items.

In this paper, for lack of space, we only briefly introduce the general notion of **multilanguage formal system** (**ML-system** from now on). Most of the paper concentrates on the representation of propositional attitudes and on a particular family of ML-systems which we prove equivalent to the modal system K and various Konolige's logics of belief [Kon84]. Konolige's work has been chosen as paradigmatic for various reasons. First, this is one of the (few) formal systems describing propositional attitudes which are easily mechanized on a computer (even considering issues like resource bounded reasoning). Second, in the authors' opinion, Konolige was the one getting closer to the notion of using multiple distinct languages even if the formal system he considered was monolanguage. Finally, some of Konolige's systems are equivalent to various modal logics (K, T, $S4$ and so on), which are the other obvious alternatives for the formalization of propositional attitudes.

The paper follows this path. In section 2 the general definition of ML-system is given and briefly discussed. Section 3 gives the basic version $MB(L, \rho)$ of the family of ML-systems considered in this paper and proves the main result of equivalence with Konolige's belief logic $B(L, \rho)$. Then, section 4 considers some particularizations of $MB(L, \rho)$ while section 5 considers an extension which deals with circumscriptive ignorance and common beliefs. Various consistency and equivalence results for all these systems are given. Finally (section 6) some concluding remarks are reported together with a brief analysis of the advantages of our system over the previous approaches.

2 ML-systems

Formally, an ML-system is defined as follows (we use natural deduction and closely follow Prawitz [Pra65] in the basic definitions and terminology; furthermore if A is a wff of a language L, we write that A is a L-wff).

Multilanguage first order theories of propositional attitudes *

Fausto Giunchiglia[1,2]
Luciano Serafini[1]
Mechanized Reasoning Group
[1]IRST, Povo, 38100 Trento, Italy
[2]DIST, University of Genoa, Via Opera Pia 11a, Genova, Italy
fausto%irst@uunet.uu.net, serafini%irst@uunet.uu.net

February 24, 1994

Long paper. Area: logics for AI, knowledge representation.

Abstract

The goal of this paper is to present a new family of formal systems, so called **multilanguage systems** (**ML-systems**), which allow the use of multiple distinct first order languages and inference rules whose premises and consequences need not belong to the same language. ML-systems are argued to formalize naturally and elegantly notions like belief, knowledge and, more in general, various forms of propositional attitudes. Some instances of ML-systems are defined and proved equivalent to the modal logic K and some of Konolige's logics for belief.

1 Introduction

Reasoning about propositional attitudes has assumed increasing importance in systems for natural language understanding, planning and knowledge representation. Most of the theoretical research on this topic has concentrated on a modal representation of propositional attitudes; some among the many examples are [Kon84, HM85], [Hal86] is a good overview, [Kon82] is one of the few counterexamples to this attitude. On the other hand a lot of the more informal work in computational linguistics (on belief ascription, speech acts, metaphor and so on) is based on the use of first order theories

*The research described here was partly supported by the Progetto Finalizzato Robotica. Richard Weyhrauch provided some of the intuitions behind the work described in this paper. It is our experience of mechanizing logic in the `FOL` system that has suggested the development of the family of formal systems introduced in this paper. Alex Simpson is also thanked. A first version of this paper was written at the Department of Computer Science of Stanford University; Caroline Talcott is thanked for providing us with all the necessary facilities.

Istituto per la Ricerca Scientifica e Tecnologica
I 38100 Trento – Loc. Pantè di Povo – tel. 0461–314444
Telex 400874 ITCRST – Telefax 0461–302040

Multilanguage First Order Theories of Propositional Attitudes

F. Giunchiglia, L. Serafini

January 1990
Technical Report # 9001-02

Published in *Proceedings SCAI-91 3rd Scandinavian Conference on Artificial Intelligence*, May 1991, pp. 228–240.

Istituto Trentino di Cultura