A logic to reason on contradictory beliefs with a majority approach

Laurence Cholvy and Christophe Garion
ONERA-Toulouse
BP 4035
31055 Toulouse Cedex 4
France
{cholvy, garion}@cert.fr

Abstract

The context of this paper is the problem of merging data provided by several information sources, which can be contradictory. It takes as a starting point, one of the majority merging operators that Konieczny and Pino-Perez have defined [KPP98], [KPP99]. In these papers, the authors have characterized, from a semantical point of view, some majority merging operators. The aim of our work is to define a logical calculus which corresponds to one of these operators. This paper focuses on the case when the information sources are sets of literals. It presents a logic which is sound and complete for some interesting kind of formulas and an associated theorem prover.

1 Introduction

The problem of merging information sources has been intensively studied for some years [BKMS91], [BKMS92], [Cho93], [Sub94], [Lin96], [LM98], [SDL+98], [KPP98], [KPP99], [Lia00]. This is due to the growing number of applications in which one needs to access several information sources to make a decision. The main problem in dealing with multiple information sources is the possible inconsistency between sources.

The many works which address this problem shows that there is not an unique method for merging information.

Obvioulsy, the adequate merging process depends on the type of the information to be merged. This information can be beliefs the sources have about the real world and in this case, the aim of the merging process is to refine our perception of the real world. But this information can also be a description of a world that is considered to be more or less ideal. This is the case for instance when merging requirements expressed by several agents about an artefact to be built or a software to be designed. In that case, the aim of the merging process is to find a consensus between the agents in order to define a description of that ideal world which agrees the agents.

But the merging process also depends on the meta-information one has about the sources. For instance, in the case of merging beliefs provided by several sources, if the respective reliability of the sources is known (in a quantitative setting or in a qualitative setting), it obviously must be used in the merging process: the more reliable a source is, the more we trust it. In the case of requirement merging, if the respective importance of the agents that provide the requirements is known, it also must be used in the merging process: the more important an agent is, the more the result of the merging must agree it.

But, if this meta-information is not known, some other types of merging processes must be defined.

Konieczny and Pino-Perez's work addresses this last case since they do not assume a priority order between the sources to be merged. For this case, they define two kinds of merging operators respectively called majority merging operators and arbitration merging operators. The first ones aim at implementing a kind of majority vote between the sources, and the second ones aim at reaching a consensus between the sources by trying to satisfy as much as possible all of them. Konieczny and Pino-Perez's work is important since it is the first time someone lists the postulates that characterize these operators (even if these postulates are expressed in the meta-language) and gives a semantical characterization of these operators.

In this present paper, we take as a starting point one of the majority merging operator that have been semantically characterized by the previous work. Our aim is to define a logic (language, model theory and proof theory) that allows us to reason with data provided by several sources and merged according to that majority operator. Indeed, even their first paper was untitled *On the logic of merging*, Konieczny and Pino-Perez did not define a logic. They focus on a semantical characterization of the merging process by defining the models of the merging result according to the models of the information sources. So, defining a logic of merging, and a calculus in particular, remained to be done.

This paper is organized as follows.

In section 2, we present a logic called MF (Majority Fusion). Its language allows us to speak about the content of the information sources (we will say what the information sources believe). Its semantic is a Kripke-type semantic and the axiomatic we give is proved to be sound and complete for some kind of formulas only in the case when the information sources are sets of atomic formulas (i.e, literals).

We prove that MF logic effectively axiomatizes a majority merging operator in section 3.

Then we present a prover, defined as a meta-program of a PROLOG-type interpretor, which allows one to automatically deduce, as a meta-theorem, the merging result. This prover is proved to be correct.

Extensions to this work are discussed in section 5.

2 The logic MF

2.1 Preliminaries

The semantics of MF logic is based on multi-sets of worlds. So, here, we recall some definitions about multi-sets.

Definition 1. A multi-set is a set where redundant occurrences are accepted. Let $MS_1 = [S_1, ..., S_n]$ and $MS_2 = [S_{n+1}, ..., S_m]$ two multi-sets. The union of two multi-sets is defined by: $MS_1 \bigsqcup MS_2 = [S_1, ..., S_m]$. The membership relation is defined by: $S \in^i MS$ iff there are exactly i occurences of S in the multi-set MS. Notice that, in the limit case, $S \in^0 MS$ iff there is no occurrence of S in MS i.e, $S \notin MS$.

Let us then introduce some notations that will be used in the rest of the paper.

Notations. If db and db' denote two information sources, then $db + db'$ will denote the information source obtained by merging db and db'. By information source, we mean any information source to be merged (in that case, we call it primitive) and also, any information source obtained after merging some information sources.

Example. For instance, if we face three (primitive) information sources db_1, db_2 and db_3 then, in particular, $db_1 + db_2$, and $(db_1 + db_2) + db_3$ are information sources but they are not primitive. The first one denotes the one obtained by merging db_1 and db_2. The second one denotes the one obtained by merging $db_1 + db_2$ and db_3.

2.2 MF language

Let us call L the language used to describe the contents of the information sources to be merged. The language L' of logic MF is obtained from L by adding several modal operators of the following form: B^i_{db} and B_{db}, where i is an integer and db denotes an information source (primitive or not).

We expect that the formula $B^i_{db}l$ means that the literal l appears exactly i times in db. And we expect that the formula $B_{db}F$ means that the information source db believes F.

Informally speaking, we introduce the modalities B^i_{db} for being able to count the occurences of a literal in an information source. The idea is that, when merging two information sources, the number of occurences of a literal is the sum of the numbers of its occurences in the two information sources respectively. Then we want that a literal is believed by an information source if the number of occurences of a literal is strictly greater than the number of occurrences of its negation.

The formal definition of L' is the following:

Definition 2. If F is a formula of L and if B^i_{db} and B_{db} are modal operators, then $B^i_{db}F$ and $B_{db}F$ are formulas of L'. If F_1 and F_2 are formulas of L' then, $\neg F_1$, $F_1 \wedge F_2$ are formulas of L'. $F_1 \vee F_2$ and $F_1 \rightarrow F_2$ are defined from the previous ones as usually.

One can notice that modal operators only govern formulas without modal operators.

Example. For instance, assume that db_1 and db_2 are the two information sources to be merged, then the modal operators we need are: $B^i_{db_1}$, B_{db_1}, $B^i_{db_2}$, B_{db_2}, $B^i_{db_1+db_2}$, and $B_{db_1+db_2}$, $B^i_{db_2+db_1}$ and $B_{db_2+db_1}$.

We expect that, for instance: $B^1_{db_1}a$ means that db_1 contains one occurrence of a. $B^0_{db_2}a$ means that db_2 contains no occurrence of a. $B^1_{db_1+db_2}a$ means that the information source obtained by merging db_1 and db_2 contains one occurrence of a. Finally, $B_{db_1+db_2}a$ means that the information source obtained by merging db_1 and db_2 believes a.

2.3 Semantics

The semantics of MF is a Kripke-type semantics [Che80]. Models are defined by:

Definition 3. Models of MF. A model of MF is a tuple $< W, val, R, B >$ such that:

- W is a set of worlds.
- val is a valuation function[1] which associates any proposition of L with a set of worlds of W.
- R is a set of functions denoted f_{db}, where db is an information source (primitive or not). Each function f_{db} associates any world of W with a multi-set of sets of worlds of W.
- B is a set of functions denoted g_{db}, where db is an information source (primitive or not). Each function g_{db} associates any world of W with a set of worlds of W.

This tuple is constrained by two constraints given below, but before, we need to give the following definition:

Definition 4. Let w and w' be two W worlds. The distance $d(w, w')$ between w and w' is defined here by the number of propositional letters p such that $w \in val(p)$ and $w' \notin val(p)$ (Hamming distance). Let $MS = [S_1...S_n]$ be a multi-set of sets of worlds. Then the distance $dsum(w, MS)$ between a world w and MS is defined by : $dsum(w, MS) = \sum_{i=1}^n Min_{w' \in S_i} d(w, w')$. Finally, any multi-set of sets of worlds MS is associated with a pre-order $\leq_{MS}$ or W, defined by: $w \leq_{MS} w'$ iff $dsum(w, MS) \leq dsum(w', MS)$.

[1] It satisfies: $val(P) \neq \emptyset$ iff P is a satisfiable propositional formula, $val(\neg P) = W \backslash val(P)$, $val(P \wedge Q) = val(P) \cap val(Q)$.

Definition 3 (continued). Models of MF.
The previous tuple $< W, val, R, B >$ is constrained by the two following constraints:

(C1) If db and db' denote two information sources, then $\forall w \in W \quad f_{db+db'}(w) = f_{db}(w) \bigsqcup f_{db'}(w)$

(C2) If db is an information source, then $\forall w \in W \quad g_{db}(w) = Min_{\leq_{f_{db}(w)}} W$

The constraint **(C1)** reflects the fact that the occurences of a litteral a in the merged information source $db + db'$ are the union of its occurences in db and of its occurrences in db'. So, it will be the case that the number of occurrences of a litteral in $db + db'$ is the sum of the number of its occurrences in db and the number of its occurrences in db',

The constraint **(C2)** corresponds, as it will be proved in section 3, to one majority merging operator defined by KPP [KPP98]. The models of the information source which is obtained by this majority merging operator, are the minimal W worlds, according to the pre-order $\leq_{f_{db}(w)}$.

Definition 5. Satisfaction of formulas.
Let $M =< W, val, R, B >$ be a model of MF and let $w \in W$. Let p be a propositional letter of L. Let F, F_1 and F_2 be formulas of L'.

$M, w \models_{MF} p$	iff	$w \in val(p)$
$M, w \models_{MF} \neg F_1$	iff	$M, w \not\models_{MF} F_1$
$M, w \models_{MF} F_1 \wedge F_2$	iff	$M, w \models_{MF} F_1$ and $M, w \models_{MF} F_2$
$M, w \models_{MF} B^i_{db} F$	iff	$val(F) \in^i f_{db}(w)$
$M, w \models_{MF} B_{db} F$	iff	$g_{db}(w) \subseteq val(F)$

Definition 6. Valid formulas in MF.
Let F be a formula of L'. F is a valid formula in MF iff $\forall M$ model of MF, $\forall w \in W$, $M, w \models_{MF} F$. We note $\models_{MF} F$.

2.4 Proof Theory

In the following, db and db' denote information sources, F and G denote formulas of L, l, l_1,..., l_n denote literals of L and i, j, k denote integers.

The axiom schemata of MF are:

$(\mathbf{A_0})$ Axiom schemata of propositional logic

$(\mathbf{A_1})$ $B_{db} \neg F \rightarrow \neg B_{db} F$

$(\mathbf{A_2})$ $B_{db} F \wedge B_{db}(F \rightarrow G) \rightarrow B_{db} G$

$(\mathbf{A_3})$ $B^i_{db} l \rightarrow \neg B^j_{db} l$ if $i \neq j$

$(\mathbf{A_4})$ $B^i_{db} l \wedge B^j_{db'} l \rightarrow B^k_{db+db'} l$ if $k = i + j$

$(\mathbf{A_5})$ $B^i_{db} l \wedge B^j_{db} \neg l \rightarrow B_{db} l$ if $i > j$

$(\mathbf{A_6})$ $B^i_{db} l \wedge B^i_{db} \neg l \rightarrow \neg B_{db} l$

$(\mathbf{A_7})$ $B_{db}(l_1 \vee ... \vee l_n) \rightarrow B_{db} l_1 \vee ... \vee B_{db} l_n$ where $\forall i \in \{1...n\}, \forall j \in \{1...n\} \;\; l_i \neq \neg l_j$

Modalities B_{db} are belief modalities and are governed by KD axioms,$(\mathbf{A_0})$, $(\mathbf{A_1})$, $(\mathbf{A_3})$.

$(\mathbf{A_3})$ says that the number of occurrences of a literal in an information source is unique.

$(\mathbf{A_4})$ express the facts that the number of occurrences of a literal in the merged information source $db + db'$ is the sum of the its occurrences in db and the number of its occurrences in db'.

$(\mathbf{A_5})$, $(\mathbf{A_6})$ express the majority aspect of the underlying merging operator. First, a literal l is believed by a source db if the number of its occurrences is strictly greater then the number of the occurrences of its negation. If the number of the occurrences of l is equal to the number of occurrences of its negation, then that literal and its negation are not believed by the information source.

$(\mathbf{A_7})$ restricts the information sources we consider to sets of literals.

The inference rules are :

(MP) If $\vdash_{MF} F$ and $\vdash_{MF} (F \rightarrow G)$ then $\vdash_{MF} G$

(Nec) $\vdash_{MF} F$ then $\vdash_{MF} B_{db} F$ for any modality B_{db}.

$\vdash_{MF} F$ denotes as usual, theorems of MF, i.e formulas that are instances of axiom schemata or that can be deduced by using axiom schemata and inference rules.

2.5 Soundness and completeness for some interesting formulas

Definition 7. Let $db_1 ... db_n$ n sets of literals to be merged, each of them being consistent. We define the formula ψ by:

$$\psi = \bigwedge_{i=1}^{n} (\bigwedge_{l \in db_i} B^1_{db_i} l \wedge \bigwedge_{l \notin db_i} B^0_{db_i} l)$$

ψ lists the information we have about the content of the given sources to be merged. More precisely, it expresses that each literal it contains has one and only one occurrence in it, and that each literal it does not contain has no occurrence in it.

The following result proves that the model theory and the proof theory previously presented are equivalent for formulas of the form $\psi \rightarrow B_{db} F$, where db is any information source.

Proposition 1. Let ψ be the formula previously defined. Let F be a formula of L and db an information source. Then we have:
$\models_{MF} \psi \rightarrow B_{db} F \iff \vdash_{MF} \psi \rightarrow B_{db} F$ and
$\models_{MF} \psi \rightarrow \neg B_{db} F \iff \vdash_{MF} \psi \rightarrow \neg B_{db} F$

Sketch of proof.

($\Longleftarrow$) In fact, we prove the soundness of the proof theory for any formula of L', by proving, as usual, that instances of the axiom schemas are valid formulas and that inference rules preserve the validity

($\Longrightarrow$) We first notice that the models $< W, val, R, B >$ in which ψ is satisfied are such that any set f_{db_i} in R is exactly the set $\{val(l) : l \in db_i\}$. Then we prove, in the same induction proof, the two implications:
$\models_{MF} \psi \rightarrow B_{db} F \implies \vdash_{MF} \psi \rightarrow B_{db} F$ *and*
$\models_{MF} \psi \rightarrow \neg B_{db} F \implies \vdash_{MF} \psi \rightarrow \neg B_{db} F$
where db is an information source (primitive or not) and F a formula of L.

Proposition 2. Let ψ be the formula previously defined. Let F be a formula of L and db an information source. Then:
(1) $\not\vdash_{MF} \psi \rightarrow B_{db}F$ *or* $\not\vdash_{MF} \psi \rightarrow \neg B_{db}F$ and
(2) $\vdash_{MF} \psi \rightarrow B_{db}F$ *or* $\vdash_{MF} \psi \rightarrow \neg B_{db}F$

Sketch of proof.

For proving (1), we notice that there is (at least) one MF-model which satisfies ψ.

For proving (2), we first notice that if M_1 and M_2 are MF-models which satisfy ψ then, $M_1 \models B_{db}F$ iff $M_2 \models B_{db}F$. Then, we also notice that if $M =< W, val, R, B >$ is a model which satisfies ψ, then $\forall w \in W \ \ \forall w' \in W \ \ g_{db}(w) = g_{db}(w')$

2.6 Example

Here, we give some examples of proofs in MF logic.

We consider three information sources: $db_1 = \{a, b\}$, $db_2 = \{a, \neg c\}$, $db_3 = \{\neg a, c\}$. By definition 7, ψ is: $B^1_{db_1}a \wedge B^1_{db_1}b \wedge B^0_{db_1}c \wedge B^0_{db_1}\neg c \wedge B^0_{db_1}\neg a \wedge B^0_{db_1}\neg b \wedge B^1_{db_2}a \wedge B^1_{db_2}\neg c \wedge B^0_{db_2}b \wedge B^0_{db_2}\neg b \wedge B^0_{db_2}\neg a \wedge B^0_{db_2}c \wedge B^1_{db_3}\neg a \wedge B^1_{db_3}c \wedge B^0_{db_3}b \wedge B^0_{db_3}\neg b \wedge B^0_{db_3}a \wedge B^0_{db_3}\neg c$

Here are some theorems of MF we can derive:

(α) $\vdash \psi \rightarrow B^2_{db_1+db_2}a$ (by (A_4))

(β) $\vdash \psi \rightarrow B^2_{(db_1+db_2)+db_3}a$ (by (α) and (A_4))

(γ) $\vdash \psi \rightarrow B^0_{db_1+db_2}\neg a$ (by (A_4))

(δ) $\vdash \psi \rightarrow B^1_{(db_1+db_2)+db_3}\neg a$ (by (γ) and (A_4))

Thus, finally, from (β), (δ) and (A_5), we can prove:

(ζ) $\vdash \psi \rightarrow B_{(db_1+db_2)+db_3}a$

This theorem means that the information source obtained by merging db_1, db_2 and db_3 believes a. Notice that this illustrates a majority attitude, since two primitive information sources believe a while only one believes $\neg a$.

In the same way, we prove:

(η) $\vdash \psi \rightarrow B_{(db_1+db_2)+db_3}b$

Thus, from (ζ), (η), (A_0) and (A_2) we prove:

$\vdash \psi \rightarrow B_{(db_1+db_2)+db_3}(a \wedge b)$

This theorem means that the information source obtained by merging db_1, db_2 and db_3 believes $(a \wedge b)$.

Similarly, we have:

$\theta : \vdash_{MF} \psi \rightarrow B^0_{db_1+db_2}c$ (by (A_4))

$\iota : \vdash_{MF} \psi \rightarrow B^1_{(db_1+db_2)+db_3}c$ (by (θ) and (A_4))

$\kappa : \vdash_{MF} \psi \rightarrow B^1_{db_1+db_2}\neg c$ (by (A_4))

$\lambda : \vdash_{MF} \psi \rightarrow B^1_{(db_1+db_2)+db_3}\neg c$ (by (θ) and (A_4))

$\mu : \vdash_{MF} \psi \rightarrow \neg B_{(db_1+db_2)+db_3}c$ (by (ι), (λ) and (A_6))

$\nu : \vdash_{MF} \psi \rightarrow \neg B_{(db_1+db_2)+db_3}\neg c$ (by (ι), (λ) and (A_6))

Thus, finally, by (μ), (ν) and (A_0) we can prove:

$\vdash_{MF} \psi \rightarrow \neg B_{(db_1+db_2)+db_3}c \wedge \neg B_{(db_1+db_2)+db_3}\neg c$

This theorem means that the information source obtained by merging db_1, db_2 and db_3 does not believe c nor $\neg c$.

3 Relation with Konieczny and Pino-Perez's work

In this section, we formally prove that MF logic allows one to reason with merged data obtained by a majority operator. More specifically, we focus on one majority merging operator defined by Konieczny and Pino-Perez. And we establish a relation between some theorems of MF and the information source obtained by that operator.

First, let us recall the definition introduced by Konieczny and Pino-Perez[2]

Let $db_1 ... db_n$ be n information sources to be merged.

Konieczny and Pino-Perez define a majority merging operator, denoted Δ_Σ, such that the models of the information source which is obtained from merging db_1 ... db_n with this operator, is semantically characterized by:

$$Mod(\Delta_\Sigma([db_1, ..., db_n])) = Min_{\leq^{\Sigma}_{[db_1 ... db_n]}}(\mathrm{W})$$

where W denotes the set of all the interpretations of the language L (the propositional language used to describe the contents of the informations sources). $\leq^{\Sigma}_{[db_1 ... db_n]}$ is a total pre-order on W defined by:

$$w \leq^{\Sigma}_{[db_1 ... db_n]} w' \ iff \ d_\Sigma(w, [db_1 ... db_n]) \leq d_\Sigma(w', [db_1 ... db_n])$$

with

$$d_\Sigma(w, [db_1 ... db_n]) = \sum_{i=1}^{n} Min_{w' \in Mod(db_i)} d(w, w')$$

where $Mod(db_i)$ is the set models of db_i and $d(w, w')$ is the Hamming distance.

In other words, when merging $db_1 ... db_n$ with the operator Δ_Σ, the result is semantically characterized by the interpretations which are minimal according to the preorder $\leq^{\Sigma}_{[db_1, ..., db_n]}$.

The following proposition establishes the relation between some theorems of logic MF and the result of this majority merging operator.

Proposition 3. Let $db_1 ... db_n$ be n sets of literals to be merged and F be a formula of L. With the notations previously introduced, we have:

$$\vdash \psi \rightarrow B_{(...(db_1+db_2)+...db_n)}F \Longleftrightarrow \Delta_\Sigma([db_1 ... db_n]) \models F$$

Sketch of proof.

We first notice that if $M =< W, val, R, B >$ is a MF-model, then for any world w in W, there is an interpretation w' of L $\{l : w' \models l\} = \{l : w \in val(l)\}$. And if w' is an interpretation of L, then there is a world w in W such that $\{l : w' \models l\} = \{l : w \in val(l)\}$. In other terms, any world in W correspond to an interpretation of L and any interpretation of L corresponds to (at least) one world in W.

[2] One will notice that we sligtly change the presentations of these definitions to remain coherent with what has already been presented.

Assume now that $\forall i = 1...n \;\; db_i = \{l_i^1 ... l_i^{m_i}\}$.

Let $M =< W, val, R, B >$ *is a* MF*-model which satifies* ψ. *Let* w *be a world in* W *and* w' *the interpretation of* L *previously characterized. We prove that:*

$$d_{sum}(w, [...val(l_1^i)...val(l_n^j)...]) = dist_\Sigma(w', [db_1...db_n])$$

This shows that the W *worlds which are minimal according to the order induced by the distance* d_{sum} *correspond to the interpretations of* L *which are minimal according to the order induced by the distance* d_Σ.

In other words, the information source whose beliefs are characterized by theorems $\vdash \psi \rightarrow B_{(...(db_1+db_2)+...db_n)}F$, is equivalent to $\Delta_\Sigma([db_1...db_n])$.

This proves that logic MF is a logic for reasoning with merged data obtained by a majority merging operator, when information sources are sets of literals.

4 Automated deduction in MF

In this section, we deal with implementation aspects. We present a theorem prover logic MF. It allows one to answer questions of the form: given the description of the information source contents, is formula F deducible after merging them? i.e, it allows one to prove theorems of the form: $\psi \rightarrow B_{db}F$.

One will notice that, in this prover, the formula ψ introduced previously, will not be used. Indeed, ψ was introduced for theoretical reasons. Its aim was to describe, in extension, what is believed and what is not believed in the primitive information sources. But in the prover, we will only need to list the explicit beliefs of the sources. Propositions which are not believed will be derived by negation as failure.

4.1 The meta-language

Let us consider a meta-language ML, based on language L, defined by:

- constants symbols of ML are propositional letters of L, names of informations sources plus a constant symbol denoted *nil* and constants denoting integers: 1, 2, etc.
- a binary function noted $*$. By convention, $(db_{i_1} * ... * db_{i_k})$ represents the term: $db_{i_1} * (db_{i_2} ... * (db_{i_k} * nil)...)$. This function will be used to denote the information sources obtained by merging information sources $db_{i_1}...db_{i_k}$.
- a binary function denoted $+$ which is the sum of integers.
- a unary function symbol $\neg$. By convention, $\neg l$ represents the term $\neg(l)$. This function will be used to describe the object-level negation.
- the binary meta-predicate symbols are: $B_{exp}, B, =$ and $>$
- A ternary meta-predicate symbol is R.
- A unary meta-predicate symbol is NIL.

The intuitive semantics of the predicates is the following:

- $B_{exp}(db, l)$ is true if literal l is explicitely stored in the primitive information source db.

- $R(db, l, i)$ is true if l appears i times in the information source db.

- $B(db, l)$ is true if the (primitive or not) information source db believes that l.

- $NIL(db)$ is true if db is *nil*.

- $i = j$ (resp, $(i > j)$) is true if integers i and j are equal (resp, if integer i is stricly greater than integer j). These two predicates will be defined in extension, in the meta-program, by a finite number of facts.

4.2 The meta-program

If there are n information sources to be merged, $db_1...db_n$, then let META be the following set of the ML formulas:

(1) $B_{exp}(db_i, l)$ if the literal l belongs to the primitive information source db_i
(2) $\neg NIL(db_2) \wedge R(db_1, l, i) \wedge R(db_2, l, j) \wedge (k = i + j) \rightarrow R(db_1 * db_2, l, k)$
(3) $NIL(db_2) \wedge B_{exp}(db_1, l) \rightarrow R(db_1 * db_2, l, 1)$[3]
(4) $NIL(db_2) \wedge \neg B_{exp}(db_1, l) \rightarrow R(db_1 * db_2, l, 0)$
(5) $R(db, l, i) \wedge R(db, \neg l, j) \wedge (i > j) \rightarrow B(db, l)$
(6) $NIL(nil)$
(7) $k = (r + l)$ and $(r + l) = k$ for any k in $\{1...n\}$ for any r in $\{1...k\}$ and for any l such that $l = k - r$
(8) $k > r$ for any k in $\{1...n\}$ and for any r in $\{1...k\}$

Notice that there is a finite number of axioms (7) and axioms (8).

The following result ensures the correctness of this meta program.

Proposition 4. Let l be a literal, let db denoting an information source (primitive or not). Then, using negation-as-failure on the meta-progran META,
(1) PROLOG succeeds in proving $B(db, l)$ if and only if $\vdash_{MF} (\psi \rightarrow B_{db} l)$;
(2) PROLOG fails if and only if $\vdash_{MF} (\psi \rightarrow \neg B_{db}\ l)$

Sketch of proof.

For proving (1), we first prove that PROLOG succeeds in proving $R(db, l, i)$ *iff* $\vdash_{MF} \psi \rightarrow B_{db}^i l$

Then, (2) derives from (1) and from Proposition 2.

This meta-program can be easily extended for proving formulas of the form: $B(db, f)$ where f is a conjunction of disjunctions of literals. This extension is trivial. We need to add to ML two new function symbols, $\vee$ and $\wedge$, for representing disjunctions of literals and conjunctions of disjunctions of literals. And also two new predicate symbols: DB and CB where: $DB(db, d)$ means that the

[3]Recall that primitive sources are sets of literals so each literal which belongs to a source has exactly one occurrence in it

information source believes the disjunction of literals d. $CB(db, c)$ means that the information source db believes the conjunction c of disjunctions of literals.

5 Conclusion

A logic for reasoning about data provided by several information source has been presented. It has been proved that the underlying merging operator it axiomatizes belongs to the class of majority merging operators defined by Konieczny and Pino-Perez. The axiomatisation has been proved to be sound and complete for some kind of interesting formulas only in the case when the information sources are sets of literals. This is a first step towards the definition of a logic of merging but this must be extended at least to the case when information sources are sets of clauses. For doing so, Lin and Mendelzon's work [LM98] provides us with a starting point since in this paper, a method for merging databases with disjonctive data is presented.

However, let us notice that, even restricted to literals, the theorem-prover developped here is powerful enough to implement a query evaluator for querying multiple relational databases where facts are ground. Such an application has already been shown, for another method of merging, in [Cho98].

Furthermore, the present work must be extended in the case when some integrity constraints are expressed. But it must be noticed that there are two ways of considering integrity constraints in the merging process. A first way is the one adopted by Konieczny and Pino-Perez and consists in constraining the merging process: the models of the merged information source are the models of the constraints which are minimal according to $\leq^{\Sigma}_{[db_1 \ldots db_n]}$. But there is another way, which consists in considering that the integrity constraints apply on the different information sources. So instead of considering the information sources one must consider new information sources obtained by integrating the constraints to the sources. The process merging remains the same.

In order to illustrate this, let us extend the example of section 2.6, and consider the constraint: $\neg(b \wedge c)$. Konieczny and Pino-Perez process characterizes $\{a, b, \neg c\}$ as result of the merging. But we claim that another way of considering this constraint is to merge the three information sources $db_i \cup IC$, i.e, $\{a, b, \neg c\}$, $\{a, \neg c\}$ and $\{\neg a, c, \neg b\}$. The majority merging then leads to $\{a, \neg c\}$.

A suggestion to extend this present work to the first case is to consider a new modal operator $\Box$ whose semantics is : $M \models \Box F$ iff $\forall w \in W \quad M, w \models F$. And to consider $\psi \wedge \Box(IC)$ instead of ψ (if IC is the set of integrity constraints). Extending this present work to the second case, seems to be easier: we just need to sligthly modify the definition of ψ in order to take the integrity constraints into account. The extension seems to be quite easy if each information source plus the integrity constraints remains equivalent to a set of literals.

References

[BKMS91] C. Baral, S. Kraus, J. Minker, and V.S. Subrahmanian. Combining multiple knowledge bases. *IEEE Trans. on Knowledge and Data Engineering*, 3(2), 1991.

[BKMS92] C. Baral, S. Kraus, J. Minker, and V.S. Subrahmanian. Combining knowledge bases consisting of first order theories. *Computational Intelligence*, 8(1), 1992.

[Che80] B. F. Chellas. *Modal logic, an introduction.* Cambridge University Press, 1980.

[Cho93] L. Cholvy. Proving theorems in a multi-sources environment. In *Proceedings of IJCAI*, pages 66–71, 1993.

[Cho98] L. Cholvy. Reasoning with data provided by federated databases. *Journal of Intelligent Information Systems*, 10(1), 1998.

[KPP98] S. Konieczny and R. Pino-Perez. On the logic of merging. In *Proc. of KR'98*, Trento, 1998.

[KPP99] S. Konieczny and R. Pino-Perez. Merging with integrity constraints. In *Proc. of ESCQARU'99*, 1999.

[Lia00] C. Liau. A conservative approach to distributed belief fusion. In *Proceedings of 3^{rd} International Conference on Information Fusion (FUSION)*, 2000.

[Lin96] J.. Lin. Integration of weighted knowldege bases. *Artificial Intelligence*, 83:363–378, 1996.

[LM98] J. Lin and A.O. Mendelzon. Merging databases under constraints. *International Journal of Cooperative Information Systems*, 7(1), 1998.

[SDL+98] S.Benferhat, D. Dubois, J. Lang, H. Prade, A. Saffiotti, and P. Smets. A general approach for inconsistency handling and merging information in prioritized knowledge bases. In *Proc. of KR'98*, Trento, 1998.

[Sub94] V.S. Subrahmanian. Amalgamating knowledge bases. *ACM Transactions on Database Systems*, 19(2):291–331, 1994.