Adaptive Arbitration by Variant Counting on Commutative Bases with Weights

Giuseppe Primiero
Centre for Logic and Philosophy of Science
Ghent University, Belgium
Email: Giuseppe.Primiero@UGent.be

Joke Meheus
Centre for Logic and Philosophy of Science
Ghent University, Belgium
Email: Joke.Meheus@UGent.be

***Abstract*—In this paper a new logical arbitration protocol for fusion of inconsistent information is designed. It defines a selection of models of a premise set in a multi-modal logic that uses the standard format of Adaptive Logics. The selected models are obtained by a counting procedure on the derivable data conflicting among the various sources. Peculiar of this approach is the definition of weights for commutative bases, in terms of the distinction between partially and fully supported information. The results obtained are compared to standard arbitration protocols and they extend previous work on the adaptive majority protocol.**

Keywords: Information Fusion, Arbitration Protocol, Adaptive Logics.

I. Introduction

Human beings deal with various forms of data and acquire knowledge by combining them. Groups of humans produce common knowledge by sharing sets of private information, but collecting complex data may require the resolution of conflicts. Similarly, processes of image fusion (“the acquisition, processing and synergistic combination of information provided by various sensors or by the same sensor in many measuring contexts”, [23]), or data fusion as the collection and classification of information clusters from heterogeneous databases (see e.g. [5], [8], [12]), are faced by the same difficulties. Both sensors and expert systems working at different times, locations, and scales, and from sources with different goals and acting in different situations can ultimately produce inconsistent information. The technology of information fusion methods can help resolve inconsistencies: probabilistic ([9]), bayesian ([15]), fuzzy ([2]) or hierarchical ([6]) models have been designed to this aim.

Data fusion allows the user to resolve redundancy, inconsistency and conflicts among data received, but errors and incompleteness are phenomena bound to occur. Data fusion architectures must not only avoid information overload as much as possible, but information loss as well. Possible architectures include centralized data fusion and processes on completely separate databases, recalled only if and when needed (see [1]). These architectures are applied in defense, intelligence and air traffic control systems and the logical problem of determining which information is preserved – and from which sources – is of a general interest for their effectiveness.

In homeland security reports, for example, to establish sensible policies on the basis of clusters of information produced by minor sources and to grant checking procedures on partially supported data might prove to be extremely important. A medical expert system providing clinical descriptions of the most likely conditions based on analyses of particular symptoms might support different solutions for the decision maker. In these and similar cases, it seems that a fusion process should account for data with different qualitative roles, avoiding to rely only on fully supported information which nonetheless might turn to be incomplete. In the present paper a logical model is introduced that explicitly accounts for fusion of inconsistent sets containing differently supported data.

In the logical literature on fusion of conflicting information, the standard distinction holding between majority and arbitration procedures accounts for different approaches to the preservation of data from various sources aiming at one consistent collection[1]. The *majority protocols* are defined by operators minimizing global dissatisfactions; operators for the *abritration protocols* aim at maximization of individual satisfaction.

An Adaptive Logic in standard format - called $\mathbf{ADM}^c$ for *Adaptive Doxastic Merging by Counting* - is formulated in [21] to model a majority procedure[2]. $\mathbf{ADM}^c$ is built up by a multi-doxastic monotonic logic, a set of formulas expressing derivable conflicts among contents and an adaptive strategy to establish which models of the premises should be selected in the fusion process. The selection of models defined by such adaptive strategy provides results corresponding to the majority merging protocol with partial support from [17], and is in turn equivalent to members of the $\triangle^{\Sigma}$ family of merging operators defined in [13]. The general structure of this logic makes it possible now to use the multi-modal framework for defining a different fusion procedure. This is obtained simply by modifying the adaptive strategy.

In the present paper the logic $\mathbf{ADM}^{c+}$ - *Adaptive Doxastic Merging by Variant Counting* - is introduced to define an arbitration protocol. In arbitration procedures, a relevant role is played by weights assigned to the bases involved in the fusion process. Such weights can be expressed as priority

[1]See [16], [17], [22] and overviews in [11], [14].

[2]For a general introduction to Adaptive Logics see [3], [4].

values (as in [10]), they can be assigned either to propositional terms (see e.g. [7]) or to the set of models of formulas (as in [22]), finally they can be formulated as possibility values (see [24]). However, the fusion process should also account for commutative bases, as enforced by postulate $A1$ for arbitration in [16] and by postulate $IC4$ in [13].

In the present formulation, the arbitration procedure relies crucially on the basic distinction between fully and partially supported bases: this distinction formulates our syntactic notion of weighted bases. We say that a literal φ is *fully supported* by some source T if $T \vDash \varphi$. A literal φ is *partially supported* by a source T if there is a set of literals Δ such that $\varphi \in \Delta$, $T \vDash \bigvee(\Delta)$, $\nvDash \bigvee(\Delta)$, and there is no $\Delta' \subset \Delta$ such that $T \vDash \bigvee(\Delta')$. As usual $\bigvee(\Delta)$ stands for the disjunction of the members of the set of literals Δ. So, where $T = \{p \wedge (q \vee r)\}$, T fully supports p but only partially supports q and r. Weights are here understood as the constraining force that the source imposes on the data produced for them to be accepted in the fusion state. The fusion protocol defined by $\mathbf{ADM}^{c+}$ satisfies both fully and partially supported data. In the case of differently weighted bases, it provides a fusion protocol that preserves their most consistent part without rejecting the basic idea of commutativity.

The paper is organised as follows. After some preliminaries in section 2, a presentation of various standard arbitration protocols is given in section 3. Then our logic is introduced in its intuitive characterization in section 4, and its semantics in section 5. Section 6 provides examples to render the comparison with the previously introduced standard protocols. Conclusions and some remarks for future research are given in the last section.

II. Preliminaries

In the following, $\mathcal{L}$ will stand for the standard language of classical propositional logic (henceforth $\mathbf{CL}$) that is formed from a finite set of atoms $\mathcal{P}$ in the usual way. A *literal* is an atom or negation of an atom; $\mathcal{P}^{\pm}$ will be used to refer to the set of literals. Metavariables for sentences of $\mathcal{L}$ will be referred to by greek letters $\varphi, \psi, \ldots$. The formulas of the language are closed under standard propositional unary and binary operators: $\neg$ (negation), $\wedge$ (conjunction), $\vee$ (disjunction), $\supset$ (implication).

A *database* T is a finite set of sentences of $\mathcal{L}$. Where $\mathcal{I}$ is a finite set of indexes $\{0, 1, \ldots, n\}$, the multi-modal language $\mathcal{L}^{\mathcal{B}}$ is $\mathcal{L}$ extended with a doxastic operator b_i, for any $i \in \mathcal{I}$. Intuitively, $b_i\varphi$ (for $i > 0$) will express that information φ is provided by information source i; the operator b_0 will be used for data collected in the fusion base as well as for the integrity constraints (if there are any): in other words for anything that results from the fusion process and for any information content to which all sources should comply. A $\mathbf{CL}$-*model* is a function $\mathcal{P} \rightarrow \{0, 1\}$. We shall use $\mathcal{M}$ to denote the set of all $\mathbf{CL}$-models. Where Γ is a set of data expressed in propositional form, a *model* M is a model of Γ iff all the members of Γ are true in it. $Mod(\Gamma)$ will denote the set of all models of Γ.

For the introduction and explanation of the arbitration protocol, a preorder on the set of models of a premise set will be used. A preorder over the set of $\mathbf{CL}$-models is a reflexive and transitive relation on $\mathcal{M}$. Where $\leq$ is a preorder, $<$ is defined as: $M < M'$ iff $M \leq M'$ and $M' \not\leq M$. Where M is a subset of $\mathcal{M}$, we say that a model M is *minimal* in M with respect to $\leq$ iff $M \in \mathsf{M}$ and there is no $M' \in \mathsf{M}$ such that $M' < M$. The expression $Min(\mathsf{M}, \leq)$ denotes the set of models that are minimal in M with respect to $\leq$.

III. Standard arbitration protocols

The basic idea of the arbitration protocols is to satisfy each information content involved in a fusion procedure, as much as this is possible in view of the required consistency. Various characterizations have been given of the arbitration procedure.

Arbitration has been introduced first in [22] as a *model-fitting* operator: this consists of a selection of the formulas from the given bases, such that the model of this selection is the one that best fits the entire knowledge base with respect to all the possible interpretations. Syntactically, this model will retain as much as possible of the original sets of formulas. The corresponding model-theoretical definition is given in terms of a preorder on the interpretations of the various bases obtained by a notion of distance $dist$.

Given two models M_1, M_2 and a database T, the preorder $M_1 \leq M_2$ holds if and only if $dist(M_1, T) \leq dist(M_2, T)$. The value of $dist$ has been variously defined, one of the most common being the *Dalal distance* from [7]. The Dalal distance $dist(M_1, M_2)$ between two models M_1 and M_2 refers to the *number of atoms* whose valuation differs in the two models. Given the set $Mod(T)$, the distance between a $\mathbf{CL}$-model M and T is then

$$dist(M, T) = min(dist(M, M')) \text{ for each } M' \in Mod(T). \quad (1)$$

The obtained model-fitting operator corresponds to an arbitration operator.

A further step in the definition of arbitration is given in [16] in terms of postulates for a commutative revision operator, one we shall denote as $\triangle^{LS}$. Its application is restricted to pairs of bases, and the commitment to the disjunctive nature of arbitration on inconsistent informations is stronger than what is usually desired, because in some cases it provides a disjunction of the bases without combinations.

Another way of looking at any such process is to focus on the formalization of the intended rationality via meta-axioms, as for example in the characterization of fusion under Integrity Constraints (IC), introduced in [13]. This is a general framework that allows the formalization of both majority and arbitration protocols. A family of arbitration operators presented in [13] is the $\triangle^{Max}$ family defined by the max function, a more fine-grained arbitration than the one obtained by $\triangle^{LS}$. To see how this family of operators works, consider a multi-set $\Psi = \{T_1, \ldots, T_n\}$ of databases; for each model M, consider the list $D = (dist_1^M, \ldots, dist_n^M)$ of distances between M and the n bases in Ψ, i.e. $dist_j^M = dist(M, T_j)$. Let L_Ψ^M be the list obtained from $(dist_1^M, \ldots, dist_n^M)$ by sorting its members in descending order. Denote now by $\leq_{lex}$ the lexicographic order among sequences of integers of the same length. For any two models M_1 and M_2, a total preorder $M_1 \leq_\Psi M_2$ holds in view of the multi-set Ψ if and only if $L_\Psi^{M_1} \leq_{lex} L_\Psi^{M_2}$. Given a multi-set Ψ and a set of integrity contraints μ, an IC-GMax operator is then defined as follows:

$$Mod(\triangle_\mu^{GMax}(\Psi)) = Min(Mod(\mu), \leq_\Psi). \quad (2)$$

The arbitration so defined works in the following way. Consider the bases:

$$T_1 = \{p \wedge q\}$$
$$T_2 = \{p \wedge q\}$$
$$T_3 = \{\neg p \wedge \neg q\}.$$

Consider now the possible models $M_1 - M_4$:

$$M_1 = \{p, q\}$$
$$M_2 = \{p, \neg q\}$$
$$M_3 = \{\neg p, q\}$$
$$M_4 = \{\neg p, \neg q\}.$$

By the definitions given above, we obtain the following lists of integers:

$$L_\Psi^{M_1} = 2, 0, 0$$
$$L_\Psi^{M_2} = 1, 1, 1$$
$$L_\Psi^{M_3} = 1, 1, 1$$
$$L_\Psi^{M_4} = 2, 2, 0.$$

The models selected by $Mod(\triangle_\mu(\Psi))$ (for $\mu = \{\emptyset\}$) are those with lower positions in the ordering of the list of distances, that is $L_\Psi^{M_2}, L_\Psi^{M_3}$, so that the result of the fusion process on databases $\Psi = \{T_1, T_2, T_3\}$ is the following:

$$\triangle^{GMax}(\Psi) = (p \wedge \neg q) \vee (\neg p \wedge q). \quad (3)$$

By setting $\mu = \{p \leftrightarrow q\}$ as an integrity contraint, it would restrict M_μ by excluding both $(p \wedge \neg q)$ and $(\neg p \wedge q)$; then the result of the fusion process would be:

$$\triangle_\mu^{GMax}(\Psi) = (p \wedge q). \quad (4)$$

By the model-fitting operator from [22], the result of the $\triangle^{GMax}(\Psi)$ operator is preserved[3], whereas by the selection designed in [16] the result of arbitration will be

$$\triangle^{LS} = (p \wedge q) \vee (\neg p \wedge \neg q). \quad (5)$$

Consider now an example where one of the bases expresses a partial support as defined in the Introduction:

$$T_1 = \{p \vee q\}$$
$$T_2 = \{\neg p\}$$
$$T_3 = \{\neg q\}.$$

[3]The model-fitting operator is in fact a specific IC-operator of the $\triangle^{Max}$ family, see [13].

The lexicographic order of distances of the bases from the standard interpretations $M_1 - M_4$ is the following:

$$L_\Psi^{M_1} = 1, 1, 0$$
$$L_\Psi^{M_2} = 1, 0, 0$$
$$L_\Psi^{M_3} = 1, 0, 0$$
$$L_\Psi^{M_4} = 1, 0, 0.$$

The interpretations with lower positions in the ranking given by the ordering are chosen, so that fusion on $\Psi = \{T_1, T_2, T_3\}$ according to arbitration gives:

$$\triangle^{GMax}(\Psi) = (p \wedge \neg q) \vee (\neg p \wedge q) \vee (\neg p \wedge \neg q). \quad (6)$$

In the following, we will introduce a fusion method based on the standard format of an Adaptive Logic that mimicks arbitration both on fully and partially supported bases.

IV. Intuitive Characterization of $\mathbf{ADM}^{c+}$

The formulation of $\mathbf{ADM}^{c+}$ is based on a multi-modal language, which enables one to represent a *set* of databases by a single set of premises[4]. It will also enable us to consider (modal) models that validate all the premises, rather than having to consider models for each of the bases separately. Remember that each different base is given an index b_i with $i \in \mathcal{I} \setminus 0$ and b_0 is used exclusively for the data holding in the fusion state, or for its integrity constraints. As it will appear from the valuation function of formulas in the next section, the b-operator is interpreted as a necessity operator. The result of the fusion procedure by the C+ strategy will be denoted as $\triangle^{c+}$ (eventually as $\triangle_\mu^{c+}$ when some set of constraints μ is given).

The bases from the last example in the previous section

$$T_1 = \{p \vee q\}$$
$$T_2 = \{\neg p\}$$
$$T_3 = \{\neg q\}$$

are given a modal translation in the set $\Gamma = \{b_1(p \vee q),\ b_2\neg p,\ \ b_3\neg q\}$.

At the basis of the adaptive logic $\mathbf{ADM}^{c+}$ is the so-called lower limit logic $\mathbf{DM}$, which is a multi-modal version of the modal logic $\mathbf{D}$.[5] In addition to all $\mathbf{CL}$-axioms and the Necessitation rule (if $\vdash_{\mathbf{CL}} \varphi$ then $\vdash_{\mathbf{DM}} b_i\varphi$), the logic $\mathbf{DM}$ validates Distribution (for any $i \in \mathcal{I}$, $b_i(\varphi \supset \psi) \supset (b_i\varphi \supset b_i\psi)$) and Consistency ($b_i\varphi \supset \neg b_i\neg\varphi$). Semantically, the $\mathbf{ADM}^{c+}$-models of a given premise set Γ are obtained by making a selection of the $\mathbf{DM}$-models of Γ. This selection is based on a special class of formulas, called *abnormalities*. In the case of the previously introduced premise set Γ, some of the $\mathbf{DM}$-models of this premise set verify the formula $b_3\neg q \supset b_0\neg q$, whereas others falsify it; or, what comes to the same, verify

[4]The multi-set Ψ consisting of n bases is introduced in [13] to the same aim.

[5]The Lower Limit Logic $\mathbf{DM}$ is the same for the adaptive logic $\mathbf{ADM}^c$ which defines the Majority protocol in [21]: this explains why the two adaptive fusion protocols belong to the same family of logics.

$b_3\neg q \wedge \neg b_0\neg q$. An abnormality is precisely a formula of the form

$$b_i\varphi \wedge \neg b_0\varphi \quad (7)$$

i.e. a formula expressing that some source i provides (full) support for a literal φ which is rejected in the fusion state.

In the case where $\Gamma = \{b_1p,\ b_2p,\ b_3\neg p\}$, there will be two types of $\mathbf{DM}$-models of Γ: those that verify b_0p and those that verify $\neg b_0p$. Models that verify b_0p, necessarily verify the abnormality $b_3\neg p \wedge \neg b_0\neg p$; those that verify $b_0\neg p$ necessarily verify the abnormalities $b_1p\wedge\neg b_0p$ and $b_2p\wedge\neg b_0p$. The selection tells us which models should be chosen.

As the arbitration protocol considers also bases that express partial support, we need a second kind of abnormalities. As far as an abnormality with respect to T_1 is concerned, one has to account for the rejection of partially supported literals. A failure to fuse the first disjunct cannot be expressed by $b_1p \wedge \neg b_0p$, because that would presuppose that p is fully supported; analogously for the second disjunct. The kind of abnormality that we need for such cases is expressed by the following form:

$$\begin{aligned}&b_i(\varphi_1 \vee \ldots \vee \varphi_n) \wedge (\neg b_i\varphi_1 \wedge \ldots \wedge \neg b_i\varphi_n)\wedge \\ &\neg b_0(\varphi_1 \vee \ldots \vee \varphi_n)\end{aligned} \quad (8)$$

where all φ_i are literals.

The desired set of models - corresponding to the result of an arbitration protocol - is obtained by selecting those $\mathbf{DM}$-models of Γ that interpret the premises "as normally as possible". Given the same lower limit logic and the same set of abnormalities, there are different ways to interpret a set of premises "as normally as possible". The meaning of this phrase is made precise by the definition of the adaptive strategy. Whereas for $\mathbf{ADM}^c$ this means to select the models that verify the minimal number of abnormalities, for $\mathbf{ADM}^{c+}$ a premise set is interpreted as normally as possible by selecting only those models that verify distinct abnormalities occurring a minimal number of times. Intuitively, the difference is the following: whereas in the case of $\mathbf{ADM}^c$ one selects, in a *global* way, the models that verify a minimal number of abnormalities (independent of their index), in the case of $\mathbf{ADM}^{c+}$ one selects those models that verify, for *each* of the indexes, a minimal number of abnormalities. The latter provides a fusion procedure that satisfies the arbitration protocol.

V. Semantic Adaptive Selection for Arbitration

The semantics of the lower limit logic $\mathbf{DM}$ of $\mathbf{ADM}^{c+}$ is a standard possible world semantics, with multiple accessibility relations. A $\mathbf{DM}$-model is a quadruple $\langle \mathcal{W}, w_o, \mathcal{R}, v\rangle$ where $\mathcal{W}$ is a set of possible worlds, $w_o \in \mathcal{W}$ is the actual world, $\mathcal{R}$ is a set of *serial*[6] accessibility relations R_i ($i \in \mathcal{I}$) over $\mathcal{W}$, and $v : \mathcal{P} \times \mathcal{W} \rightarrow \{0, 1\}$ is an assignment function.

[6]That accessibility relations are serial is needed to ensure that $b_i\varphi \supset \neg b_i\neg\varphi$ is valid.

The valuation function defined by a model M is characterized as follows:

C1 where $A \in \mathcal{P}$, $v_M(A, w) = v(A, w)$
C2 $v_M(\neg A, w) = 1$ iff $v_M(A, w) = 0$
C3 $v_M(A \vee B, w) = 1$ iff $v_M(A, w) = 1$ or $v_M(B, w) = 1$
C4 $v_M(A \wedge B, w) = 1$ iff $v_M(A, w) = 1$ and $v_M(B, w) = 1$
C5 $v_M(A \supset B, w) = 1$ iff $v_M(A, w) = 0$ or $v_M(B, w) = 1$
C6 $v_M(b_i\varphi, w) = 1$ iff $v_M(\varphi, w') = 1$ for all w' such that R_iww'

The standard semantic notions are defined as usual: a model M verifies A iff $v_M(A, w_0) = 1$, $\Gamma \models_{\mathbf{DM}} A$ iff all $\mathbf{DM}$-models of Γ verify A, and $\models_{\mathbf{DM}} A$ iff all $\mathbf{DM}$-models verify A.

The set of abnormalities consists of the union of the sets of the formulas defined in equations 7 and 8:

***Definition** 1 (Set of Abnormalities Ω):* $\Omega = \{b_i\varphi \wedge \neg b_0\varphi \mid i \in \mathcal{I}\backslash 0, \varphi \in \mathcal{P}^{\pm}\}\cup\{b_i(\varphi_1\vee\ldots\vee\varphi_n)\wedge(\neg b_i\varphi_1\wedge\ldots\wedge\neg b_i\varphi_n)\wedge \neg b_0(\varphi_1 \vee \ldots \vee \varphi_n) \mid i \in \mathcal{I} \setminus 0, \varphi_1, \ldots, \varphi_n \in \mathcal{P}^{\pm}, n > 1\}$.

This set contains all abnormal formulas for fully and partially supported literals.

The selection is based on the mentioned crucial role of the abnormalities: in $\mathbf{ADM}^{c+}$, a disjunction of abnormalities may be $\mathbf{DM}$-derivable without any of its disjuncts being $\mathbf{DM}$-derivable. Consider as an example the simple premise set $\Gamma = \{b_1p, b_2\neg p\}$: from it, neither $b_1p\wedge\neg b_0p$ nor $b_2\neg p\wedge\neg b_0\neg p$ is $\mathbf{DM}$-derivable, but their disjunction $(b_1p\wedge\neg b_0p)\vee(b_2\neg p\wedge \neg b_0\neg p)$ is. It is in view of this property that the *adaptive strategy* is needed: its aim is to specify what it means to interpret a premise set as "normally as possible" in the case where one or more disjunctions of abnormalities are derivable.

In order to define the semantic selection of $\mathbf{ADM}^{c+}$, some further definitions are needed, in line with the standard format of adaptive logics. Disjunctions of abnormalities will be called *Dab*-formulas, and the abbreviation $Dab(\Delta)$ will be used to refer to them:

***Definition** 2 (Dab-Formula):* $Dab(\Delta)$ stands for $\bigvee(\Delta)$ where $\Delta \subseteq \Omega$.

If Δ is a singleton, $Dab(\Delta)$ is a single abnormality; if $\Delta = \emptyset$, any disjunction $A \vee Dab(\Delta)$ corresponds to A. A *Dab*-formula that is $\mathbf{DM}$-derivable from Γ will be called a *Dab*-consequence of Γ:

***Definition** 3 (Dab-Consequence):* $Dab(\Delta)$ is a *Dab*-consequence of a premise set Γ iff $\Gamma \models_{\mathbf{DM}} Dab(\Delta)$.

If $Dab(\Delta)$ is a *Dab*-consequence of a set Γ, then so is any $Dab(\Delta')$ such that $\Delta' \supset \Delta$. This is why a further definition is needed:

***Definition** 4 (Minimal Dab-Consequence):* A disjunction of abnormalities $Dab(\Delta)$ is a *minimal* *Dab*-consequence of Γ iff $\Gamma \models_{\mathbf{DM}} Dab(\Delta)$ and there is no $\Delta' \subset \Delta$ such that

$\Gamma \models_{\mathbf{DM}} Dab(\Delta')$.

The selection by variant Counting will consider the various disjunctions of abnormalities verified by the **DM**-models of a premise set Γ; in particular, it will focus on the number of abnormalities verified in any given model involving a certain index. For any given model of the premise set, the selection will register the number of data provided by each source and that cannot be selected in the fusion state according to that model. Among these models, it will choose those that have the minimal number of such disagremeents, according to some ordering.

Consider first all formulas $A \in \Omega$ such that the b-operator indexed 1 occurs in A: typically, this will be the set of all formulas that are of the form $b_1\phi \wedge \neg b_0\phi$ or of the form $b_1(\phi \vee \psi) \wedge (\neg b_1\phi \wedge \neg b_1\psi) \wedge \neg b_0(\phi \vee \psi))$. Call this set Ω^1. Then consider the set of all formulas of the same kind occurring with b-operator indexed 2 and call this set Ω^2, and so on up to index n.

The set Ω from Definition 1 is in turn the union of all the various Ω^i sets:

***Definition** 5 (The Set of all indexed Abnormalities):*

$$\Omega = \bigcup_{i=1}^{n} \Omega^i$$

It is obvious that one can consider now the set of abnormalities with a given index Ω^i as a proper subset of Ω.

For each model of a given premise set, we now consider the abnormal formulas of a certain Ω^i verified by that model:

***Definition** 6 (The abnormal part of a model with index i):* $Ab^i(M) = \{A \mid A \in \Omega^i \text{ and } M \models A\}$.

For any model M_j of a given premise set, let $\mathcal{C}^i_{M_j} = |Ab^i(M_j)|$ denote the cardinality of its abnormal part with respect to Ω^i:

***Definition** 7 (Abnormal cardinality of a model):* Given a model M_j of a premise set Γ and its abnormal part $Ab^i(M_j)$, its abnormal cardinality $\mathcal{C}^i_{M_j}$ is the number of abnormal formulas $A \in \Omega^i$ verified in the model M_j.

For each model M, we construct the list $(\mathcal{C}^1_M, \dots, \mathcal{C}^n_M)$, where n is the number of elements of $\mathcal{I}$. Let L^M_Γ be the list obtained from $(\mathcal{C}^1_M, \dots, \mathcal{C}^n_M)$, by sorting its elements in descending order. Let now $\leq_{lex}$ be the lexicographic order between sequences of integers of the same length. On the basis of the ordering $\leq_{lex}$, a total preorder $\leq^{\mathcal{C}}_\Gamma$ holds among the various models $M_1, \dots, M_n$ of Γ in the following way:

***Definition** 8 (Selection of Models by $\mathbf{ADM}^{c+}$):* A total preorder $\leq^{\mathcal{C}}_\Gamma$ holds between models of a premise set Γ according to the following definition

$$M_i \leq^{\mathcal{C}}_\Gamma M_j \text{ iff } L^{M_i}_\Gamma \leq_{lex} L^{M_j}_\Gamma. \quad (9)$$

Where M_Γ stands for the set of **DM**-models of Γ, the variant Counting strategy $\triangle^{c+}(\Gamma)$ will select among those models the minimal ones with respect to the ordering obtained by $\leq^{\mathcal{C}}_\Gamma$:

$$Mod(\triangle^{c+}(\Gamma)) = Min(\mathsf{M}_\Gamma,\ \leq^{\mathcal{C}}_\Gamma). \quad (10)$$

VI. Some Examples.

Let us consider a few examples to show how the variant Counting selection works in practice. Consider the premise set:

$$\Gamma = \{b_1(p \vee q), b_2\neg p, b_3\neg q\} \quad (11)$$

where each $b_i\varphi$ expresses the content reported by the information source i. Let us derive the minimal disjunction of abnormalities from Γ ($!b_i\varphi$ will abbreviate $b_i\varphi \wedge \neg b_0\varphi$ provided $\varphi \in \mathcal{P}^\pm$ and $!b_i(\varphi_1 \vee \dots \vee \varphi_n)$ will abbreviate $b_i(\varphi_1 \vee \dots \vee \varphi_n) \wedge (\neg b_i\varphi_1 \wedge \dots \wedge \neg b_i\varphi_n) \wedge \neg b_0(\varphi_1 \vee \dots \vee \varphi_n)$ provided each $\varphi_i \in \mathcal{P}^\pm$ and $n > 1$):

$$Dab(\Delta) = Dab(\{!b_1(p \vee q),\ !b_1p,\ !b_1q,\ !b_2\neg p,\ !b_3\neg q\}). \quad (12)$$

This Dab-consequence of Γ provides the following subsets of $\Omega^1 - \Omega^3$:

$$\omega^1 = \{!b_1(p \vee q), !b_1p, !b_1q\}$$
$$\omega^2 = \{!b_2\neg p\}$$
$$\omega^3 = \{!b_3\neg q\}.$$

The various models of Γ can be defined according to the data verified in the fusion state, with the following obvious alternatives:

$$M_1 = \{b_0p, b_0q\}$$
$$M_2 = \{b_0p, b_0\neg q\}$$
$$M_3 = \{b_0\neg p, b_0q\}$$
$$M_4 = \{b_0\neg p, b_0\neg q\}.$$

For any of these models a certain number of abnormalities will be verified. Correspondingly, for each model M_j and any indexed set of abnormalities Ω^i there will be a value to $\mathcal{C}^i_{M_j}$. These values are listed as follows – we assume that M_1-M_4 verify a minimal number of abnormalities for each of the indexes:

	Ω^1	Ω^2	Ω^3
M_1	0	1	1
M_2	0	1	0
M_3	0	0	1
M_4	1	0	0

At the intersection of each M_j and Ω^i, one has the value of $\mathcal{C}^i_{M_j}$, so that $|Ab^1(M_1)| = 0$, $|Ab^2(M_1)| = 1$, $|Ab^3(M_1)| = 1$ and the ordering of these values is $1, 1, 0$.

The total preorder $\leq^{\mathcal{C}}_\Gamma$ among the various models is obtained on the basis of $\leq_{lex}$ (in the following, models with equal position in the ordering have been adjoined as different indexes to the same model symbol):

$$M_{2,3,4} \leq^{\mathcal{C}}_{\Gamma} M_1. \tag{13}$$

This means in turn that the result of $Mod(\triangle^{c+}(\Gamma))$ will select the three alternatives that are equally listed as members of $Min(\mathsf{M}, \leq^{\mathcal{C}}_{\Gamma})$:

$$\triangle^{c+}(\Gamma) = b_0((p \wedge \neg q) \vee (\neg p \wedge q) \vee (\neg p \wedge \neg q)). \tag{14}$$

This result is the intuitive one for an arbitration process preserving non conflicting data and providing the consistent combination of the highest number possible of literals.

Another example adapted from [13] will complete the presentation of $\mathbf{ADM}^{c+}$ with respect to the generalization to a premise set containing only bases with full support. Consider the following databases listed as a premise set:

$$\begin{aligned}\Gamma = \{&b_1(p \wedge q \wedge r), b_2(p \wedge q \wedge r),\\ &b_3(\neg p \wedge \neg q \wedge \neg r \wedge \neg s), b_4(q \wedge r)\}\end{aligned} \tag{15}$$

and let $\mu = \{((p \wedge q) \vee (p \wedge r) \vee (q \wedge r)) \supset s\}$ be the set of integrity constraints imposed on the fusion state[7]. Our usual procedure will provide the following subsets of Ω^1-Ω^4:

$$\begin{aligned}\omega^1 &= \{!b_1 p, !b_1 q, !b_1 r\}\\ \omega^2 &= \{!b_2 p, !b_2 q, !b_2 r\}\\ \omega^3 &= \{!b_3 \neg p, !b_3 \neg q, !b_3 \neg r\}\\ \omega^4 &= \{!b_4 q, !b_4 r\}.\end{aligned}$$

Let us now consider our models, with respect to the literals involved in the abnormalities:

$$\begin{aligned}M_1 &= \{b_0 p, b_0 q, b_0 r\}\\ M_2 &= \{b_0 p, b_0 q, b_0 \neg r\}\\ M_3 &= \{b_0 p, b_0 \neg q, b_0 r\}\\ M_4 &= \{b_0 p, b_0 \neg q, b_0 \neg r\}\\ M_5 &= \{b_0 \neg p, b_0 q, b_0 r\}\\ M_6 &= \{b_0 \neg p, b_0 q, b_0 \neg r\}\\ M_7 &= \{b_0 \neg p, b_0 \neg q, b_0 r\}\\ M_8 &= \{b_0 \neg p, b_0 \neg q, b_0 \neg r\}.\end{aligned}$$

In view of the models $M_1 - M_8$ for the literals p, q, r that are conflicting in Γ, the values of the (minimal) abnormal cardinalities are the following:

	Ω^1	Ω^2	Ω^3	Ω^4
M_1	0	0	3	0
M_2	1	1	2	1
M_3	1	1	2	1
M_4	2	2	1	2
M_5	1	1	2	0
M_6	2	2	1	1
M_7	2	2	1	1
M_8	3	3	0	2

[7] According to the premise set, the literal s is explicitly denied in some of the bases, whereas none of them explicitly accept it. According to standard arbitration, in particular by postulate $A2$ in [16] and postulate $IC5$ in [13], this implies the direct selection of $\neg s$. The meaning of the integrity constraints μ is to avoid those combinations of literals that imply s.

At the intersection of each M_j and Ω^i, one has the value of $\mathcal{C}^i_{M_j}$; the corresponding values are $|Ab^1(M_1)| = 0$, $|Ab^2(M_1)| = 0$, $|Ab^3(M_1)| = 3$, $|Ab^4(M_1)| = 0$ and so on for the other models. The ordering of these cardinalities is the following:

$$\begin{aligned}&\leq^{M_1}_{lex} = 3, 0, 0, 0\\ &\leq^{M_2}_{lex} = 2, 1, 1, 1\\ &\leq^{M_3}_{lex} = 2, 1, 1, 1\\ &\leq^{M_4}_{lex} = 2, 2, 2, 1\\ &\leq^{M_5}_{lex} = 2, 1, 1, 0\\ &\leq^{M_6}_{lex} = 2, 2, 1, 1\\ &\leq^{M_7}_{lex} = 2, 2, 1, 1\\ &\leq^{M_8}_{lex} = 3, 3, 2, 0.\end{aligned}$$

The obtained total preorder has therefore the following structure:

$$M_5 \leq^{\mathcal{C}}_{\Gamma} M_{2,3} \leq^{\mathcal{C}}_{\Gamma} M_{6,7} \leq^{\mathcal{C}}_{\Gamma} M_4 \leq^{\mathcal{C}}_{\Gamma} M_1 \leq^{\mathcal{C}}_{\Gamma} M_8. \tag{16}$$

According to the imposed constraints μ, the models that verify $(p \wedge q)$ or $(p \wedge r)$ or $(q \wedge r)$ are avoided: this is the case for model M_5 and models M_2 and M_3. The result of the fusion process according to $Min(\mathsf{M}_\Gamma, \leq^{\mathcal{C}}_{\Gamma})$ is therefore given by models $M_{6,7}$:

$$\triangle^{c+}_{\mu}(\Gamma) = b_0((\neg p \wedge \neg q \wedge r \wedge \neg s) \vee (\neg p \wedge q \wedge \neg r \wedge \neg s)). \tag{17}$$

As it is shown by this last example, also in the case of bases expressing full support the semantic selection performed by $\mathbf{ADM}^{c+}$ provides the standard result of the $\triangle^{GMax}$ family of operators from [13].

VII. Conclusion

The selection of models of a premise set by $\mathbf{ADM}^{c+}$ provides results corresponding to those obtained for standard arbitration in [22] and in [13] for the operators of the $\triangle^{GMax}$ family. The adaptive interpretation of arbitration allows moreover to formalize an intuitive system of weights for databases on the basis of the distinction between partially and fully supported information. This distinction, combined with the result of the arbitration protocol for data fusion, is crucial for the selection of information in those contexts where partial support can influence decision making processes in an effective way.

The adaptive framework has an important unifying aspect with respect to the various logical protocols of information fusion. As mentioned in the introduction, by the same multi-modal logic defined in this paper and the distinct strategy of simple Counting defined in [21], one obtains a semantic selection equivalent to majority for bases with partial support in [17] and to the $\triangle^{\Sigma}$ family from [13]. Therefore, the two strategies design logics belonging to the very same family. Next steps of this research are directed to the definition of other fusion protocols. In particular, in [20] a different strategy within the family of logics $\mathbf{ADM}$ let to obtain a quasi-merging protocol equivalent to the $\triangle^{Max}$ family of operators from [13]. A further advantage of the adaptive format for merging, is that a syntactic formulation of our operators is feasible in view of the proof-theory of the standard format of Adaptive Logics.

Applications of the various adaptive fusion strategies are to be considered for standard problematic decision making procedures. One immediate example is represented by the formulation of a solution to the Discursive Dilemma (also known as the Doctrinal Paradox) by using the selection procedure designed for the majority protocol (see [19]). It is very relevant to the understanding of the various adaptive selections that the majority protocol is able to provide an effective solution only by a variation obtained by relaxing the universal domain condition, whereas on the other hand the selection by Variant Counting seems to provide a different but very effective solution to the mentioned dilemma in a straightforward way. This seems to support the arguments that show limitations related to data fusion protocols based on a majority rule.

Acknowledgements

Research for this paper was supported by subventions from Ghent University and from the Research Foundation – Flanders (FWO - Vlaanderen).

References

[1] Akita, R.M. (2002) "User based data fusion approaches", Proceedings of the Fifth International Conference on Information Fusion, 2002, vol.2, pp. 1457-1462, DOI 10.1109/ICIF.2002.1020987.

[2] Bao-Hua Yang, Ping He, Ben-Li Wang (2003), "Research on fuzzy model and algorithms for data fusion", International Conference on Machine Learning and Cybernetics, 2003, Volume 4, Issue , 2-5 Nov. 2003 Page(s): 2480 - 2484.

[3] Batens, D. (2001), "A general Characterization of Adaptive Logics", *Logique & Analyse*, 173-175, pp. 45-68.

[4] Batens, D. (2007), "A Universal Logic Approach to Adaptive Logics", *Logica universalis*, 1 (2007), pp. 221-242, Birkhäuser Verlag Basel/Switzerland.

[5] Batini, C., Lenzerini, M. (1986) "A comparative analysis of methodologies for database schema integration", ACM Computing Surveys 18 (4), pp.323-364, 1986.

[6] Carpenter, G.A., Martens, S., Ogas, O.J. (2005) "Self-organizing information fusion and hierarchical knowledge discovery: a new framework using ARTMAP neural networks", *Neural Networks* vol.8. issue 13, pp.287-295.

[7] Dalal, M. (1988), "Investigations into a theory of knowledge base revision: Preliminary Report", in Proceedings of the Seventh National Conference on Artificial Intelligence (AAAI-88), 1988, pp. 475-479.

[8] Elmagarmid, A., Rusinliewicz, M., Sheth, A. (1999) *Management of Heterogeneous and Autonomous Database Systems*, Morgan Kaufmann, 1999.

[9] Feili Yu, Levchuck, G., Pattipati, K., Fang Tu (2007), "A Probabilistic computational model for identifying organizational structures from uncertain message data", Proceedings of the 10th International Conference on Information Fusion, 2007, DOI 10.1109/ICIF.2007.4407970.

[10] Fagin, R., Ullman, J.D., Vardi, M.Y. (1983), "On the Semantics of Updates in Databases", in Proceedings of Second ACM SIGACT-SIGMOD, Atlanta, pp.352-365.

[11] Grégoire, E., Konieczny, S. (2006), "Logic-based Approaches to Information Fusion", *Information Fusion*, 7 (2006), pp.4-18.

[12] Kim, W. ed. (1995), *Modern database Systems: The Object Model, Interoperability and Beyond*, Addison Wesley, 1995.

[13] Konieczny, S., Pino-Pérez, R. (2002), "Merging Information under Constraints: A Logical Framework", *Journal of Logic and Computation*, vol.12 n. 5, pp. 773-808.

[14] Konieczny, S., Pino-Pérez, R. (2002), "On the Frontier between Arbitration and Majority", Proceedings of the 8^{th} International Conference on Principles of Knowledge Representation and Reasoning, 2002, pp.109-120.

[15] Kumar, M., Garg, D.P., Zachery, R.A. (2006), "A generalized approach for inconsistency detection in data fusion from multiple sensors", *American Control Conference 2006*.

[16] Liberatore, P., Schaerf, M. (1998), "Arbitration (or how to merge knowledge bases)", IEEE Transactions on Knowledge and Data Engineering, 10 (1), pp. 76-90.

[17] Lin, J., Mendelzon, A.O. (1999), "Knowledge base merging by majority", in *Dynamic Worlds: from the Frame Problem to Knowledge Management*, Kluwer, Dordrecht, pp. 195-218.

[18] Primiero, G. (2007), "Belief Merging based on Adaptive Interaction", in *A Meeting of the Minds*, Proceedings of the Workshop on Logic, Rationality and Interaction (LORI) 2007, Beijing, China, Texts in Computer Science, vol.8, College Publications, London, pp. 315-320.

[19] Primiero, G., "Solving the discursive dilemma by selecting disagreements", presented at LOFT2008, 8^{th} Conference on Logic and the Foundations of Game and Decision Theory, 2008.

[20] Primiero, G., Meheus, J. (2008), "Quasi-merging and Pure-arbitration on Information for the Family of Adaptive Logics $\mathbf{ADM}$", Proceedings of the Workshop on Logic and Intelligent Interaction, ESSLLI2008.

[21] Primiero, G., Meheus, J. (2008), "Majority Merging by Adaptive Counting", forthcoming in a Special Issue of *Knowledge, Rationality and Action (Synthese)*.

[22] Revesz, P.Z. (1997) "On the Semantics of Arbitration", *Journal of Algebra and Computation*, 7(2), pp. 133-160.

[23] Simone, G., Farina, A., Morabito, F.C. , Serpico, S.B. and Bruzzone, L. (2002) "Image fusion techniques for remote sensing applications", *Information Fusion*, 3 (2002), pp. 3-15.

[24] Zadeh, L.A. (1978) "Fuzzy sets as a basis for a theory of possibility", *Fuzzy Sets and Systems*, vol.1, no.1, pp.3-28.