On Merging Strategy-Proofness

Patricia Everaere and **Sébastien Konieczny** and **Pierre Marquis**
Centre de Recherche en Informatique de Lens (CRIL-CNRS)
Université d'Artois - F-62307 Lens - France
{everaere, konieczny, marquis}@cril.univ-artois.fr

Abstract

Merging operators aim at defining the beliefs/goals of a group of agents from the beliefs/goals of each member of the group. Whenever an agent of the group has preferences over the possible results of the merging process (i.e. the possible merged bases), she can try to rig the merging process by lying on her true beliefs/goals if this leads to a better merged base according to her point of view. Obviously, strategy-proof operators are highly desirable in order to guarantee a fair merging process even when some of them are not sincere. In fact, when strategy-proofness is not guaranteed, it may be questioned whether the result of the merging process actually represents the beliefs/goals of the group. In this paper, the strategy-proof landscape for many merging operators from the literature, including model-based ones and formula-based ones, is drawn. Both the general case and several restrictions on the merging process (among others, the number of agents and the presence of integrity constraints), are considered.

Introduction

Merging operators aim at defining the beliefs/goals of a group of agents from the beliefs/goals of each member of the group. Though beliefs and goals are distinct notions, merging operators can typically be used for merging either beliefs or goals. Thus, most of the logical properties given in (Konieczny & Pino Pérez 1998; 1999) for characterizing rational belief merging operators can be used for characterizing as well rational goal merging operators.

Whatever beliefs or goals are merged, there are numerous situations where agents have preferences on the possible results of the merging process (i.e. the merged bases). As far as goals are concerned, an agent is surely satisfied when her individual goals are chosen as goals of the group. In the case of belief merging, an agent can be interested in imposing her beliefs to the group (i.e. "convincing" the other agents), especially because the result of a further decision stage may depend on the beliefs of the group.

So, as soon as an agent participates to a merging process, the *strategy-proofness* problem has to be considered. The question is: is it possible for a given agent to improve the result of the merging process with respect to her own point of view by lying on her true beliefs/goals, given that she knows the beliefs/goals of each agent of the group and the way beliefs/goals are merged. When strategy-proofness is not guaranteed, it may be questioned whether the result of the merging process actually represents the beliefs/goals of the group.

For an example of manipulation, let us consider the following illustrating scenario:

Example 1 *Marie, Alain and Pierre always spend their evening together. They have to plan what they will do this evening. Pierre wants to go to a restaurant for diner, but not to the movie. Marie does not want to go out for diner. Alain does not want to stay at home, i.e. he wants to go to the restaurant or to the movie. If one uses a usual merging operator for defining the goal of the group*[1]*, then the goal of the group will be to go out for diner and not to the movie. Marie will not be very happy… However, if Marie lies and claims that she wants to go to the movie but not to the restaurant, then the result of the merging process will be different. Indeed, in this case, the goal of the group will be to go either to the restaurant or to the movie: Marie may still avoid to go out for diner.*

The aim of this work is to draw the strategy-proofness landscape for many merging operators from the literature, including model-based ones and formula-based ones. For each operator under consideration, we aim at determining whether it is strategy-proof in the general case, and under some restrictions on the merging process (including the number of agents and the presence of integrity constraints) and on the set of available strategies for the agents.

The rest of the paper is organized as follows. First we give the needed basic definitions. We then recall the definitions of the main propositional belief merging operators of the literature. After what, we give several definitions of strategy-proofness based on a general notion of satisfaction index. Then, we report our strategy-proofness results. We end by noticing related work, just before a concluding discussion. Due to space limitations, only the most typical (and short) proofs are given in the paper.

[1]Formally, the model-based operator $\triangle^{d,\Sigma}$.

Preliminaries

We consider a propositional language $\mathcal{L}$ defined in the standard way from a finite set of propositional variables $\mathcal{P}$ and the usual connectives, including $\top$ and $\bot$.

An interpretation (or world) is a total function from $\mathcal{P}$ to $\{0, 1\}$, denoted by a bit vector whenever a strict total order on $\mathcal{P}$ is specified. The set of all interpretations is noted $\mathcal{W}$. An interpretation ω is a model of a formula $\phi \in \mathcal{L}$ if and only if it makes it true in the usual truth functional way. $[\phi]$ denotes the set of models of formula ϕ, i.e. $[\phi] = \{\omega \in \mathcal{W} \mid \omega \models \phi\}$.

A belief/goal base K denotes the set of beliefs/goals of an agent, it is a finite and consistent set of propositional formulas, interpreted conjunctively. $\bigwedge K$ denotes the singleton belief/goal base containing the conjunction of every formula of K. We say that a base K is complete if it has a unique model, such a base is noted K_ω.

A belief/goal profile E denotes the group of agents that is involved in the merging process. It is a multi-set (bag) of belief/goal bases $E = \{K_1, \ldots, K_n\}$ (hence two agents are allowed to exhibit identical bases).

$\subseteq$ will denote set containment and $\subset$ strict set containment, i.e. $A \subset B$ if and only if $A \subseteq B$ and $A \neq B$. The multi-set union is noted $\sqcup$ and the multi-set containment relation is noted $\sqsubseteq$. The cardinal of a finite set A is noted $\#(A)$. The same notation is used for a multi-set (the cardinal of a finite multi-set is the sum of the numbers of occurrences of each of its elements). We note by $\bigwedge E$ the conjunction of bases of E, i.e. $\bigwedge E = \bigwedge K_1 \wedge \ldots \wedge \bigwedge K_n$, where $\bigwedge K_i$ denotes the conjunction of all formulas from K_i ($i \in 1 \ldots n$). A profile E is said to be consistent if and only if $\bigwedge E$ is consistent.

A pre-order $\leq$ is a reflexive and transitive relation. A pre-order is total if $\forall \omega, \omega'$ $\omega \leq \omega'$ or $\omega' \leq \omega$. Let $\leq$ be a pre-order, we define the corresponding strict ordering $<$ as $\omega < \omega'$ if and only if $\omega \leq \omega'$ and $\omega' \not\leq \omega$, and the induced equivalence relation (indifference) $\simeq$ is given by $\omega \simeq \omega'$ if and only if $\omega \leq \omega'$ and $\omega' \leq \omega$. We write $\omega \in min(A, \leq)$ if and only if $\omega \in A$ and $\nexists \omega' \in A$ s.t. $\omega' < \omega$.

The result of the merging of the bases of a profile E, under the integrity constraints μ is the base denoted $\triangle_\mu(E)$. The integrity constraints consist of a consistent formula the merged base has to satisfy (it may represent some physical laws, some norms, etc.).

Merging operators

We recall in this section the two main families of merging operators from the literature. The first family is defined by a selection of some interpretations, usually using a notion of distance. The second family is defined by a selection of some formulas in the set-theoretic union of the bases. For more details on those two families, see for example (Konieczny, Lang, & Marquis 2002; 2004).

Model-based operators

The first family is based on the selection of some interpretations, the "closest" ones to the given profile (Revesz 1997; Konieczny & Pino Pérez 1998; 1999; Lin & Mendelzon 1999; Liberatore & Schaerf 1998).

Definition 1 *A* pseudo-distance *between interpretations is a total function $d : \mathcal{W} \times \mathcal{W} \mapsto \mathbb{R}^+$ such that for any ω, ω', $\omega'' \in \mathcal{W}$:*
- $d(\omega, \omega') = d(\omega', \omega)$*, and*
- $d(\omega, \omega') = 0$ *if and only if* $\omega = \omega'$.

A distance *between interpretations is a pseudo-distance that satisfies triangular inequality:*
- $d(\omega, \omega') \leq d(\omega, \omega'') + d(\omega'', \omega')$.

Two widely used distances between interpretations are Dalal distance (Dalal 1988), denoted d_H, which is the Hamming distance between interpretations (i.e. the number of propositional variables on which the two interpretations differ); and the drastic distance, denoted d_D, which is the simplest pseudo-distance one can define: it gives 0 if the two interpretations are the same one, and 1 otherwise.

Definition 2 *An* aggregation function f *is a total function associating a nonnegative real number to every finite tuple of nonnegative real numbers and s.t. for any $x_1, \ldots, x_n, x, y \in \mathbb{R}^+$:*
- *if* $x \leq y$*, then* $f(x_1, \ldots, x, \ldots, x_n) \leq f(x_1, \ldots, y, \ldots, x_n)$. *(non-decreasingness)*
- $f(x_1, \ldots, x_n) = 0$ *if and only if* $x_1 = \ldots = x_n = 0$. *(minimality)*
- $f(x) = x$. *(identity)*

Widely used functions are the $\max$ (Revesz 1997; Konieczny & Pino Pérez 2002b), the sum Σ (Revesz 1997; Lin & Mendelzon 1999; Konieczny & Pino Pérez 1999), or the leximax $GMax$ (Konieczny & Pino Pérez 1999; 2002b).

The chosen distance between interpretations induces a "distance" between an interpretation and a base, which in turn gives a "distance" between an interpretation and a profile, using the aggregation function. This latter distance gives the needed notion of closeness $\leq_E$ (a pre-order induced by E).

Definition 3 *Let d be a pseudo-distance between interpretations and f be an aggregation function. The result $\triangle_\mu^{d,f}(E)$ of the (model-based) merging of E given the integrity constraints μ is defined by:*
- $d(\omega, K) = min_{\omega' \models K} d(\omega, \omega')$.
- $d(\omega, E) = f_{\{K_i \in E\}}(d(\omega, K_i))$.
- $\omega \leq_E \omega'$ *if and only if* $d(\omega, E) \leq d(\omega', E)$.
- $[\triangle_\mu^{d,f}(E)] = \min([\mu], \leq_E)$.

Let us step back to the example given in introduction in order to illustrate the way model-based merging operators work:

Example 2 *Consider the set $\mathcal{P}$ with two propositional variables m(ovie) and r(estaurant), taken in this order. The goals of the three agents are then given by the following bases: K_1 whose set of models is $\{00, 10\}$ (Marie's wishes), K_2 whose set of models is $\{01, 10, 11\}$ (Alain's wishes) and*

$[\mu]$	$d_H(\omega, K_1)$	$d_H(\omega, K_2)$	$d_H(\omega, K_3)$	$\Delta_\mu^{d_H,\Sigma}(\{K_1, K_2, K_3\})$
00	0	1	1	2
01	1	0	0	1
10	0	0	2	2
11	1	0	1	2

Table 1: Merging with $\Delta_\mu^{d_H,\Sigma}$.

K_3 whose set of models is $\{01\}$ (Pierre's wishes). There is no integrity constraint ($\mu = \top$).

Table 1 sums up the computations of the result of the merging for the operator $\Delta_\mu^{d_H,\Sigma}$. It states, for each model of the integrity constraints (first column) the distances between this model and the bases, and in the last column the aggregate distance between the model and the profile. According to this operator, $[\Delta_\mu^{d_H,\Sigma}(\{K_1, K_2, K_3\}] = \{01\}$, so the goal of the group is to go to the restaurant and not to the movie.

Formula-based merging

The other main family of merging operators is composed of operators usually called "formula-based operators" or "syntax-based operators", since the syntactic form of the bases at play may easily influence the result of the merging process: replacing a base $\{\varphi_1, \ldots, \varphi_n\}$ by the (logically equivalent) base $\{\varphi_1 \wedge \ldots \wedge \varphi_n\}$ may lead to change the corresponding merged base (while it is not the case for model-based operators). Formula-based operators are based on the selection of consistent subsets of formulas in the union of the bases of the profile E. Several operators are obtained by letting vary the selection criterion. The result of the merging process is the set of consequences that can be inferred from all selected subsets. See (Baral *et al.* 1992; Baral, Kraus, & Minker 1991; Rescher & Manor 1970; Konieczny 2000) for more details.

Definition 4 $\textsc{maxcons}(K, \mu)$ *is the set of all M that satisfy:*

- *$M \subseteq K \cup \{\mu\}$, and*
- *$\mu \in M$, and*
- *If $M \subset M' \subseteq K \cup \{\mu\}$, then M' is not consistent.*

When maximality must be taken w.r.t. cardinality, we use the notation $\textsc{maxcons}_{card}(K, \mu)$. To be more precise, $\textsc{maxcons}_{card}(K, \mu)$ is the set of all M that satisfy:

- *$M \subseteq K \cup \{\mu\}$, and*
- *$\mu \in M$, and*
- *If $\#(M) < \#(M')$ with $\{\mu\} \subseteq M' \subseteq K \cup \{\mu\}$, then M' is not consistent.*

Let $\textsc{maxcons}(E, \mu) = \textsc{maxcons}(\bigcup_{K_i \in E} K_i, \mu)$.

The following operators have been defined so far (Baral, Kraus, & Minker 1991; Baral *et al.* 1992; Konieczny 2000):

Definition 5 *Let E be a profile and let μ be an integrity constraint:*

$\triangle_\mu^{C1}(E) = \bigvee\{M \in \textsc{maxcons}(E, \mu)\}$.
$\triangle_\mu^{C3}(E) = \bigvee\{M \mid M \in \textsc{maxcons}(E, \top)$ *and* $M \cup \{\mu\}$ *consistent*$\}$.
$\triangle_\mu^{C4}(E) = \bigvee\{M \in \textsc{maxcons}_{card}(E, \mu)\}$.
$\triangle_\mu^{C5}(E) = \bigvee\{M \cup \{\mu\} \mid M \in \textsc{maxcons}(E, \top)$ *and* $M \cup \{\mu\}$ *consistent*$\}$
if this set is not empty and μ otherwise.

$\triangle_\mu^{C1}(E)$, $\triangle_\mu^{C3}(E)$ and $\triangle_\mu^{C4}(E)$ correspond respectively to $Comb1(E, \mu)$, $Comb3(E, \mu)$ and $Comb4(E, \mu)$ as defined in (Baral *et al.* 1992). There is no $\triangle^{C2}$ operator since the corresponding $Comb2$ operator is equivalent to $Comb1$ (Baral *et al.* 1992). The $\triangle^{C5}$ operator is a slight modification of $\triangle^{C3}$ in order to get more logical properties (Konieczny 2000). Note that unlike the other operators, $\triangle_\mu^{C3}$ may generate inconsistent merged bases.

Let us step back to the example given in introduction:

Example 3 *Marie, Alain and Pierre wishes can be respectively encoded by the following bases $K_1 = \{\neg r\}$, $K_2 = \{m \vee r\}$ and $K_3 = \{\neg m, r\}$, so $E = \{K_1, K_2, K_3\}$. The maximal (w.r.t. set inclusion) consistent subsets in the union of those bases are (there is no integrity constraint $\mu = \top$): $\{\neg r, m \vee r\}$, $\{\neg r, \neg m\}$, and $\{m \vee r, \neg m, r\}$. So for this example we get $\triangle_\mu^{C1}(E) = \triangle_\mu^{C3}(E) = \triangle_\mu^{C5}(E) \equiv \neg r \vee \neg m$, that means that the goal of the group is not to go to the movie or not to go to the restaurant. And with $\triangle_\mu^{C4}(E) \equiv \neg m \wedge r$, that means that the group has to go to the restaurant but not to the movie.*

Note also that if we encode Pierre's wishes as $K_3' = \{\neg m \wedge r\}$, so with $E' = \{K_1, K_2, K_3'\}$, then the maximal (w.r.t. set inclusion) consistent subsets in the union of those bases are $\{\neg r, m \vee r\}$ and $\{m \vee r, \neg m \wedge r\}$. So, in this case, we get $\triangle_\mu^{C1}(E') = \triangle_\mu^{C3}(E') = \triangle_\mu^{C4}(E') = \triangle_\mu^{C5}(E') \equiv (m \wedge \neg r) \vee (\neg m \wedge r)$. Accordingly, the goal of the group is to go either to the movie or to the restaurant, but not to both.

This last example shows that syntax-sensitivity can be an additional mean of manipulation. For example, if one uses the $\triangle^{C4}$ merging operator, then it is more interesting for Pierre to express his wishes trough K_3 rather than with K_3' (although those two bases are logically equivalent). In the first case the result of the merging is logically equivalent to K_3, it is no more the case with K_3' (see (Lafage & Lang 2000) for a similar example of syntax-sensitivity in a preference aggregation context).

Other valuable formula-based operators can be defined, for instance by replacing each base K by the singleton containing the conjunction of its elements before making the

$[\mu]$	$d_H(\omega, K_1)$	$d_H(\omega, K_1')$	$d_H(\omega, K_2)$	$d_H(\omega, K_3)$	$\Delta_\mu^{d_H,\Sigma}(\{K_1, K_2, K_3\})$	$\Delta_\mu^{d_H,\Sigma}(\{K_1', K_2, K_3\})$
00	0	1	1	1	2	3
01	1	2	0	0	1	2
10	0	0	0	2	2	2
11	1	1	0	1	2	2

Table 2: $\Delta_\mu^{d_H,\Sigma}$ is not strategy-proof for i_{d_w}.

union. As the comma - which is a specific connective, even if it is not truth-functional - is usually not equivalent to standard conjunction in the formula-based framework, such operators may easily give merged bases that differ from their original counterpart, as illustrated in the above example. Clearly enough, the resulting operators are not any longer sensitive to the syntactic presentation of the bases (replacing every base by a logically equivalent one leads to the same merged base). Formally, we have:

Definition 6 *Let $E = \{K_1, \ldots, K_n\}$ be a profile and let μ be an integrity constraint:*

- $\triangle_\mu^{\widehat{C1}}(E) = \triangle_\mu^{C1}(\{\bigwedge K_1, \ldots, \bigwedge K_n\})$.
- $\triangle_\mu^{\widehat{C3}}(E) = \triangle_\mu^{C3}(\{\bigwedge K_1, \ldots, \bigwedge K_n\})$.
- $\triangle_\mu^{\widehat{C4}}(E) = \triangle_\mu^{C4}(\{\bigwedge K_1, \ldots, \bigwedge K_n\})$.
- $\triangle_\mu^{\widehat{C5}}(E) = \triangle_\mu^{C5}(\{\bigwedge K_1, \ldots, \bigwedge K_n\})$.

See (Konieczny 2000; Konieczny, Lang, & Marquis 2002; 2004) for other refinements of formula-based operators, that allow a finer use of the distribution of the information.

Strategy-proofness

The strategy-proofness issue for a merging operator can be stated as follows: is it possible for a given agent to improve the result of the merging process with respect to her own point of view by lying on her true beliefs/goals, given that she knows the beliefs/goals of each agent of the group and the way beliefs/goals are merged? If this question can be answered positively, then the operator is not strategy-proof (the agent may benefit from being untruthful). Thus, a merging operator is not strategy-proof if one can find a profile $E = \{K_1, \ldots, K_n\}$ which represents the bases of the other agents, an integrity constraint μ, and two bases K and K' such that the result of the merging of E and K' is better for the agent than the result of the merging of E with her true base K.

Definition 7 (strategy-proofness)
Let i be a satisfaction index, i.e. a total function from $\mathcal{L} \times \mathcal{L}$ to $\mathbb{R}$. A merging operator Δ is strategy-proof *for i if and only if there is no integrity constraint μ, profile $E = \{K_1, \ldots, K_n\}$, base K and base K' s.t.*

$$i(K, \Delta_\mu(E \sqcup \{K'\})) > i(K, \Delta_\mu(E \sqcup \{K\})).$$

A profile E is said to be manipulable *by a base K for index i given a merging operator Δ and an integrity constraint μ if and only if there exists a base K' s.t.*

$$i(K, \Delta_\mu(E \sqcup \{K'\})) > i(K, \Delta_\mu(E \sqcup \{K\})).$$

Clearly, there are numerous different ways to define the satisfaction of an agent given a merged base. Many *ad hoc* definitions can be considered. This is closely related to the problem of measuring how similar two logical bases are, hence it is close to verrisimilitude issues (see e.g. (Kuipers 1987)).

The following three indexes are meaningful when no additional information are available.

The first two indexes are drastic ones: they range to $\{0, 1\}$, so the agent is either fully satisfied or not satisfied at all.

Definition 8 (weak drastic index)

$$i_{d_w}(K, K_\Delta) = \begin{cases} 1 & \textit{if } \bigwedge K \wedge \bigwedge K_\Delta \textit{ is consistent,} \\ 0 & \textit{otherwise.} \end{cases}$$

This index takes value 1 if the result of the merging (noted K_Δ in the definition) is consistent with the agent's base K, and 0 otherwise. It means that the agent is considered fully satisfied as soon as its beliefs/goals are consistent with the merged base.

Definition 9 (strong drastic index)

$$i_{d_s}(K, K_\Delta) = \begin{cases} 1 & \textit{if } K_\Delta \models K, \\ 0 & \textit{otherwise.} \end{cases}$$

This index takes value 1 if the agent's base is a logical consequence of the result of the merging, and 0 otherwise. In order to be fully satisfied, the agent must impose her beliefs/goals to the whole group.

The last index is not a boolean one, leading to a more gradual notion of satisfaction. The more compatible the merged base with the agent's base the more satisfied the agent. The compatibility degree of K with K_Δ is the (normalized) number of models of K that are models of K_Δ as well:

Definition 10 (probabilistic index)

$$i_p(K, K_\Delta) = \frac{\#([K] \cap [K_\Delta])}{\#([K_\Delta])}.$$

When $\#([K_\Delta]) = 0$, we set $i_p(K, K_\Delta) = 0$.

$i_p(K, K_\Delta)$ is the probability to get a model of K from a uniform sampling in the models of K_Δ. This index takes its minimal value when no model of K is in the models of the merged base K_Δ, and its maximal value when each model of the merged base is a model of K. Strategy-proofness for these three indexes are not independent notions:

$[\mu]$	$d_H(\omega, K_1)$	$d_H(\omega, K_1')$	$d_H(\omega, K_2)$	$\Delta_\mu^{d_H,\Sigma}(\{K_1, K_2\})$	$\Delta_\mu^{d_H,\Sigma}(\{K_1', K_2\})$
01	0	0	2	2	2
10	1	2	0	1	2
11	1	1	1	2	2

Table 3: Manipulability of $\Delta^{d_H,\Sigma}$ for i_{d_w} and i_p for two bases.

Proposition 1

- *If a merging operator is strategy-proof for i_p, then it is strategy-proof for i_{d_w}.*
- *Consider a merging operator that generates only consistent bases. If it is strategy-proof for i_p, then it is strategy-proof for i_{d_s}.*

On the other hand, one can prove that strategy-proofness for i_{d_w} and strategy-proofness for i_{d_s} are logically independent in the general case (an operator can be strategy-proof for one of them without being strategy-proof for the other, and it can be strategy-proof for both of them or for neither).

Let us conclude this section with our running example, and give formal arguments explaining how Marie can manipulate the merging process:

Example 4 *We consider three bases, $[K_1] = \{00, 10\}$ (Marie's wishes), $[K_2] = \{01, 10, 11\}$ (Alain's wishes) and $[K_3] = \{01\}$ (Pierre's wishes). There is no constraint ($\mu = \top$). $[\Delta_\mu^{d_H,\Sigma}(\{K_1, K_2, K_3\})] = \{01\}$ and $i_{d_w}(K_1, \Delta_\mu^{d_H,\Sigma}(\{K_1,\ K_2, K_3\})) = 0$, which means that Marie is not satisfied. If Marie reports K_1' whose set of models is $\{10\}$ instead of K_1, then $[\Delta_\mu^{\Sigma}(\{K_1', K_2, K_3\})] = \{01, 10, 11\}$ and $i_{d_w}(K_1, \Delta_\mu^{d_H,\Sigma}(\{K_1', K_2, K_3\})) = 1$, that is more satisfactory from her point of view. Table 2 gives the details of the computations for this example.*

Strategy-proofness results

In the general case, both the family of model-based operators and the family of formula-based operators are not strategy-proof for the three indexes we consider. This means that there are operators from those families which are not strategy-proof.

However, imposing further restrictions may lead to strategy-proofness. Considering them in a systematic way allows us to draw the strategy-proofness landscape for both families.

A first restriction concerns the number of bases to be merged. The interesting case is when $\#(E) = 2$, where we can sometimes reach strategy-proofness whereas for larger profiles the operator is manipulable. Since $\{\top\}$ typically plays the role of a neutral element for all the operators we consider, in the sense that for every E, μ, we have $\triangle_\mu(E) \equiv \triangle_\mu(E \sqcup \{\top\})$, manipulation is monotonic with respect to the number of bases for those operators: if an operator is manipulable for $\#(E) = n$, then it is manipulable for $\#(E) > n$.

A second parameter is the completeness of the beliefs/goals of the agent who aims at manipulating. In some cases, having such strong beliefs/goals renders any strategy impossible.

A third significant parameter is the presence of integrity constraints. On the one hand, adding nontrivial integrity constraints ($\mu \not\equiv \top$) can render a manipulation possible, while it is not when no integrity constraints are considered. The other way, adding integrity constraints may prevent from any manipulation (simply by choosing a constraint μ that is not consistent with the base K of the untruthful agent) which would be possible otherwise.

Another restriction bears on the possible available strategies. In the general case the untruthful agent is free from reporting any base, even if it is "quite far" from her true base. However, there are numerous situations for which the other agents participating to the merging process have some information about her true base. In the following, we consider two restrictions on available strategies (and the corresponding notions of strategy-proofness): the erosion (resp. dilatation) manipulation is when the reported base K' is necessarily logically stronger (resp. weaker) than the true one K. The erosion (resp. dilatation) manipulation is safe for the untruthful agent when the other agents may only have access to a subset of the countermodels (resp. models) of her true beliefs/goals.

Model-based operators

The main strategy-proofness result for model-based merging operators in the general case holds when the drastic distance d_D is considered:

Proposition 2 *Let f be any aggregation function. $\Delta^{d_D,f}$ is strategy-proof for i_p, i_{d_w} and i_{d_s}.*

As shown on the running example, the family obtained by considering the Hamming distance is not strategy-proof. Let us now focus on this family, and consider successively the two operators obtained by considering Σ and $GMax$ as aggregation functions.

As to $\Delta^{d_H,\Sigma}$, the number of bases and the presence of integrity constraints are significant. Let's see first that $\Delta^{d_H,\Sigma}$ is not strategy-proof in the general case.

Proposition 3 *$\Delta^{d_H,\Sigma}$ is not strategy-proof for any of i_{d_w}, i_{d_s} and i_p, even if there are only two bases involved in the merging process.*

Proof : The following example shows the manipulability of $\Delta^{d_H,\Sigma}$ for i_{d_w} (and then for i_p). Let us consider the constraint $\mu = a \vee b$ and the two bases K_1 and K_2 defined by their respective sets of models: $[K_1] = \{00, 01\}$ and $[K_2] = \{10\}$. We have $[\Delta_\mu^{d_H,\Sigma}(\{K_1, K_2\})] = \{10\}$ and

$[\mu]$	$d_H(\omega, K_1)$	$d_H(\omega, K_1')$	$d_H(\omega, K_2)$	$\Delta_\mu^{d_H,\Sigma}(\{K_1, K_2\})$	$\Delta_\mu^{d_H,\Sigma}(\{K_1', K_2\})$
100	1	2	1	2	3
110	1	1	2	3	3
111	0	0	2	2	2

Table 4: Manipulability of $\Delta^{d_H,\Sigma}$ for i_{d_s} for two bases.

$i_{d_w}(K_1, \Delta_\mu^{d_H,\Sigma}(\{K_1, K_2\})) = 0$. On the other hand, if the agent whose base is K_1 gives K_1', with $[K_1'] = \{01\}$ instead of K_1, we obtain $[\Delta_\mu^{d_H,\Sigma}(\{K_1', K_2\})] = \{01, 10, 11\}$ and $i_{d_w}(K_1, \Delta_\mu^{d_H,\Sigma}(\{K_1', K_2\}) = 1$. This example shows the manipulability of $\Delta^{d_H,\Sigma}$ even if there are only two bases in the profile. The computations are detailed in Table 3.

The following example shows the manipulability of $\Delta^{d_H,\Sigma}$ for i_{d_s}. Let us consider the constraint $\mu = (a \wedge b) \vee (a \wedge \neg b \wedge \neg c)$ and the two bases K_1 and K_2 defined by their sets of models: $[K_1] = \{000, 111\}$ and $[K_2] = \{000, 001\}$. We have $[\Delta_\mu^{d_H,\Sigma}(\{K_1, K_2\})] = \{111, 100\}$ and $i_{d_s}(K_1, \Delta_\mu^{d_H,\Sigma}(\{K_1, K_2\})) = 0$. On the other hand, if the agent whose base is K_1 gives K_1', with $[K_1'] = \{111\}$ instead of K_1, we obtain $[\Delta_\mu^{d_H,\Sigma}(\{K_1', K_2\})] = \{111\}$ and $i_{d_s}(K_1, \Delta_\mu^{d_H,\Sigma}(\{K_1', K_2\})) = 1$. This example shows the manipulability of $\Delta^{d_H,\Sigma}$ for i_{d_s} even if there are only two belief bases in the profile. The computations are detailed in Table 4.

□

When no integrity constraints are considered (i.e. $\mu \equiv \top$), any operator $\Delta_\top^{d,\Sigma}$ (where d is a distance) is strategy-proof for the indexes i_{d_w} and i_{d_s} when only two bases are considered:

Proposition 4 *Let d be any distance. Provided that only two bases are to be merged, $\Delta_\top^{d,\Sigma}$ is strategy-proof for the indexes i_{d_w} and i_{d_s}.*

But this result holds only for two bases since we have the following:

Proposition 5 *$\Delta_\top^{d_H,\Sigma}$ is not strategy-proof for indexes i_{d_w} and i_{d_s} if at least three bases are involved in the merging process.*

Imposing further constraints may protect from any manipulation:

Proposition 6 *For any distance d, $\Delta^{d,\Sigma}$ is strategy-proof for the indexes i_p, i_{d_w} and i_{d_s} when the initial base K is complete.*

Proof :

- i_{d_w} and i_{d_s}. The property is a direct consequence of Proposition 12, showing that if $\Delta^{d,\Sigma}$ is manipulable for i_{d_w} and i_{d_s} by a belief base K, then it is manipulable by erosion. But this is impossible whenever K is complete.
- i_p. *Reductio ad absurdum*: let us suppose that there is an operator $\Delta^{d,\Sigma}$, where d is any distance, which is manipulable for i_p given a complete base $[K_{\omega_1}] = \{\omega_1\}$. So, there exists an integrity constraint μ, a profile E, and a base K' s.t.:

  $$i_p(K_{\omega_1}, \Delta_\mu^{d,\Sigma}(K_{\omega_1} \sqcup E)) < i_p(K_{\omega_1}, \Delta_\mu^{d,\Sigma}(\{K'\} \sqcup E)).$$

  If we have $i_p(K_{\omega_1}, \Delta_\mu^{d,\Sigma}(K_{\omega_1} \sqcup E)) = 0$, then we have $i_{d_w}(K_{\omega_1}, \Delta_\mu^{d,\Sigma}(K_{\omega_1} \sqcup E)) = 0$ too. In that case, manipulation for i_p implies manipulation for i_{d_w} but we proved that no manipulation is possible for i_{d_w}. As a consequence, we can suppose that $i_p(K_{\omega_1}, \Delta_\mu^{d,\Sigma}(K_{\omega_1} \sqcup E)) \neq 0$. Consequently:

  $$\frac{\#(K_{\omega_1} \cap [E \triangle_\mu^\Sigma K_{\omega_1}])}{\#([E \triangle_\mu^\Sigma K_{\omega_1}])} \neq 0$$

  (where $E \triangle_\mu^\Sigma K_{\omega_1}$ is a light notation for $\Delta_\mu^{d,\Sigma}(K_{\omega_1} \sqcup E)$). This equation allows us to infer that ω_1 is a model of $E \triangle_\mu^\Sigma K_{\omega_1}$. In order to increase $i_p(K_{\omega_1}, \Delta_\mu^{d,\Sigma}(K' \sqcup E))$, we have to reduce the number of models of $E \triangle_\mu^\Sigma K'$ compared to $E \triangle_\mu^\Sigma K_{\omega_1}$, without removing ω_1 from $[E \triangle_\mu^\Sigma K']$. So we have to find $\omega_2 \neq \omega_1$ s.t.:

  $$\omega_2 \models E \triangle_\mu^\Sigma K_{\omega_1} \text{ and } \omega_2 \not\models E \triangle_\mu^\Sigma K'.$$

  So, $\omega_2 \models \mu$, and we have:

  $$d(\omega_2, E \sqcup K_{\omega_1}) = d(\omega_1, E \sqcup K_{\omega_1}) \quad (1)$$

  and:

  $$d(\omega_2, E \sqcup \{K'\}) > d(\omega_1, E \sqcup \{K'\}) \quad (2)$$

  (because ω_1 is a model of both $E \triangle_\mu^\Sigma K_{\omega_1}$ and $E \triangle_\mu^\Sigma K'$). With the aggregation function Σ, we get from equation 1:

  $$d(\omega_2, \omega_1) + d(\omega_2, E) = d(\omega_1, E)$$

  and from equation 2:

  $$d(\omega_2, K') + d(\omega_2, E) > d(\omega_1, K') + d(\omega_1, E).$$

  Replacing $d(\omega_1, E)$ by $d(\omega_2, \omega_1) + d(\omega_2, E)$, we obtain:

  $$d(\omega_2, K') + d(\omega_2, E) > d(\omega_1, K') + d(\omega_2, \omega_1) + d(\omega_2, E),$$

  so:

  $$d(\omega_2, K') > d(\omega_1, K') + d(\omega_2, \omega_1).$$

  If ω_1' is a model of K' s.t. $d(\omega_1, K') = d(\omega_1, \omega_1')$, then we have:

  $$d(\omega_2, K') > d(\omega_1, \omega_1') + d(\omega_2, \omega_1).$$

  Furthermore, by definition of min, we have $d(\omega_2, \omega_1') \geq d(\omega_2, K')$, so:

  $$d(\omega_2, \omega_1') > d(\omega_1, \omega_1') + d(\omega_2, \omega_1)$$

  which contradicts the triangular inequality.

$[\mu]$	$d_H(\omega, K_1)$	$d_H(\omega, K'_1)$	$d_H(\omega, K_2)$	$\Delta_\mu^{d_H,GMax}(\{K_1, K_2\})$	$\Delta_\mu^{d_H,GMax}(\{K'_1, K_2\})$
000	1	0	3	$(3,1)$	$(3,0)$
001	0	1	2	$(2,0)$	$(2,1)$
010	2	1	2	$(2,2)$	$(2,1)$
011	1	2	1	$(1,1)$	$(2,1)$
100	2	1	2	$(2,2)$	$(2,1)$
101	1	2	1	$(1,1)$	$(2,1)$
110	3	2	1	$(3,1)$	$(2,1)$
111	2	3	0	$(2,0)$	$(3,0)$

Table 5: Manipulability of $\Delta^{d_H,GMax}$ for i_{d_w} with two complete bases.

$\square$

In contrast to $\Delta^{d_H,\Sigma}$, $\Delta^{d_H,GMax}$ is not strategy-proof even in very restricted situations:

Proposition 7

- *$\Delta^{d_H,GMax}$ is not strategy-proof for the satisfaction indexes i_{d_w} and i_p, even if there is no constraint ($\mu \equiv \top$), the initial base K is complete and only two agents are involved in the merging process.*
- *$\Delta^{d_H,GMax}$ is strategy-proof for i_{d_s} when:*
  - *two bases are considered, and*
  - *$\mu \equiv \top$, and*
  - *the initial base is complete.*

  If one of these conditions is not satisfied, then $\Delta^{d_H,GMax}$ is no more strategy-proof for id_s.

Proof : We just give the proof for i_{d_w} and i_p. Table 5 shows the manipulability of $\Delta^{d_H,GMax}$ for the weak satisfaction index i_{d_w} and two complete bases. We consider K_1 s.t. $[K_1] = \{001\}$ and K_2 with $[K_2] = \{111\}$, and $\mu = \top$. We have $[\Delta_\mu^{d_H,GMax}(\{K_1, K_2\})] = \{011, 101\}$, so no model of K_1 belongs to $\Delta_\mu^{d_H,GMax}(\{K_1, K_2\})$ and $i_{d_w}(K_1, \Delta_\mu^{d_H,GMax}(\{K_1, K_2\}) = 0$. If agent 1 gives K'_1 with $[K'_1] = \{000\}$ instead of K_1, then $[\Delta_\mu^{d_H,GMax}\{K'_1, K_2\})] = \{001, 010, 011, 100, 101, 110\}$ and $i_{d_w}(K_1, \Delta_\mu^{d_H,GMax}(\{K'_1, K_2\}) = 1$. Since manipulability for i_{d_w} holds, manipulability for i_p holds as well.

$\square$

Formula-based operators

For the probabilistic index, almost none of the formula-based operators under consideration is strategy-proof:

Proposition 8

- *No operator among $\triangle^{C1}$, $\triangle^{C3}$, $\triangle^{C4}$, $\triangle^{C5}$ is strategy-proof for i_p, even if there are only two agents involved in the merging process, there is no constraint ($\mu \equiv \top$) and the initial base K is complete.*
- *$\triangle^{\widehat{C1}}$ and $\triangle^{\widehat{C5}}$ are strategy-proof for i_p if there are only two agents involved in the merging process, but none of them is strategy-proof for i_p if three agents or more are involved in the merging process, even if $\mu \equiv \top$ and if the initial base is complete.*
- *$\triangle^{\widehat{C3}}$ is strategy-proof for i_p if there are only two agents involved in the merging process and if $\mu \equiv \top$, but it is not strategy-proof for two agents if $\mu \not\equiv \top$, or if three agents or more are involved in the merging process, even if the initial base is complete.*
- *$\triangle^{\widehat{C4}}$ is strategy-proof for i_p.*

Proof : We only give here an example of manipulation of $\triangle^C$ for i_p, with $\#(E) = 2$, a complete base K, and $\mu = \top$. Consider $E = \{K_1, K_2\}$, with $K_1 = \{a \wedge b\}$ and $K_2 = \{\neg(a \wedge b)\}$. Then $\triangle_\top^{C1}(E) \equiv \top$, and $i_p(K_1, \triangle_\top^{C1}(E)) = \frac{1}{4}$. But if agent 1 gives $K'_1 = \{a, b\}$ instead of K_1, then $\triangle_\top^{C1}(\{K'_1, K_2\}) \equiv a \vee b$, and $i_p(K_1, \triangle_\top^{C1}(\{K'_1, K_2\})) = \frac{1}{3}$. So E is manipulable by K_1 for i_p. The same example holds for $\triangle^{C4}$. It remains to note now that $\triangle_\top^{C1} = \triangle_\top^{C3} = \triangle_\top^{C5}$ to conclude the first point of the proof.

$\square$

For the two drastic indexes, the results are more nuanced:

Proposition 9

- *$\triangle^{C1}$ and $\triangle^{\widehat{C1}}$ are strategy-proof for both i_{d_w} and i_{d_s}.*
- *$\triangle^{C3}$ and $\triangle^{\widehat{C3}}$ are not strategy-proof for any of i_{d_w}, i_{d_s}, even if there are only two bases involved in the merging process and the initial base K is complete, but they are strategy-proof for both indexes if $\mu = \top$.*
- *$\triangle^{C4}$ is not strategy-proof for any of i_{d_w}, i_{d_s}, even if there are only two bases involved in the merging process, the initial base K is complete, and $\mu \equiv \top$.*
- *$\triangle^{\widehat{C4}}$ is strategy-proof for i_{d_w} and i_{d_s}.*
- *$\triangle^{C5}$ and $\triangle^{\widehat{C5}}$ are not strategy-proof for both i_{d_w} and i_{d_s} in the general case. But strategy-proofness can be achieved in the following restricted cases: $\triangle^{C5}$ and $\triangle^{\widehat{C5}}$ are strategy-proof for both i_{d_w} and i_{d_s} if $\mu \equiv \top$. $\triangle^{C5}$ and $\triangle^{\widehat{C5}}$ are strategy-proof for i_{d_w} if the initial base is complete. $\triangle^{\widehat{C5}}$ is strategy-proof for both i_{d_w} and i_{d_s} if there are only two bases involved in the merging process.*

Ensuring strategy-proofness: the case of complete bases

Let us now focus on a very specific case: the situation where every base is complete. While this situation is rather unfrequent when dealing with belief bases to be merged, it can be imposed in a goal merging setting, especially if it guarantees strategy-proofness. This explains why we consider such a case in this paper.

Proposition 10 *The strategy-proofness landscape for several merging operators under the restriction each base is complete is given in Table 6. f is any aggregation function, d is any distance, $\surd$ means "strategy-proof", $-$ means "non strategy-proof" even if $\#(E) = 2$ and $\mu \equiv \top$, $-^*$ means "non strategy-proof" even if either $\#(E) = 2$ or $\mu \equiv \top$, but "strategy-proof" if both $\#(E) = 2$ and $\mu \equiv \top$. Finally, $-^\top$ means "non strategy-proof" even if $\#(E) = 2$, but "strategy-proof" whenever $\mu \equiv \top$.*

Δ	i_p	i_{d_w}	i_{d_s}
$\Delta^{d_D,f}$	$\surd$	$\surd$	$\surd$
$\Delta^{d,\Sigma}$	$\surd$	$\surd$	$\surd$
$\Delta^{d_H,GMax}$	$-$	$-$	$-^*$
$\triangle^{C1}$	$-$	$\surd$	$\surd$
$\triangle^{C3}$	$-$	$-^\top$	$-^\top$
$\triangle^{C4}$	$-$	$-$	$-$
$\triangle^{C5}$	$-$	$\surd$	$-^\top$
$\triangle^{\widehat{C1}}$	$\surd$	$\surd$	$\surd$
$\triangle^{\widehat{C3}}$	$\surd$	$\surd$	$\surd$
$\triangle^{\widehat{C4}}$	$\surd$	$\surd$	$\surd$
$\triangle^{\widehat{C5}}$	$\surd$	$\surd$	$\surd$

Table 6: Strategy-proofness of merging operators when each base is complete.

As Proposition 10 shows, no operator among $\Delta^{d_H,GMax}$ and the $\triangle^C$ ones ensures full strategy-proofness in the restricted case where two complete bases are to be merged and no integrity constraint is considered. Contrastingly, all the other operators offer strategy-proofness for the three indexes whenever every base is complete.

Restricted strategies

We will now focus on two restrictions on the available strategies for the untruthful agents. The erosion (resp. dilatation) manipulation is when the reported base K' is necessarily logically stronger (resp. weaker) than the true one K.

Definition 11 *Let i be a satisfaction index.*

- *A merging operator is* erosion strategy-proof *for i if and only if there is there is no integrity constraint μ, profile $E = \{K_1, \ldots, K_n\}$, base K and base K' s.t. $K' \models K$ and $i(K, \Delta_\mu(E \sqcup \{K'\})) > i(K, \Delta_\mu(E \sqcup \{K\}))$.*
- *A merging operator is* dilatation strategy-proof *for i if and only if there is there is no integrity constraint μ, profile $E = \{K_1, \ldots, K_n\}$, base K and base K' s.t. $K \models K'$ and $i(K, \Delta_\mu(E \sqcup \{K'\})) > i(K, \Delta_\mu(E \sqcup \{K\}))$.*

The erosion (resp. dilatation) manipulation is safe for the untruthful agent when the other agents may only have access to a subset of the countermodels (resp. models) of her true beliefs/goals.

The first result gives the dilatation strategy-proofness of model-based operators:

Proposition 11 *Let d be a pseudo-distance and let f be an aggregation function. $\Delta^{d,f}$ is dilatation strategy-proof for the indexes i_p, i_{d_w} and i_{d_s}.*

This result has to be compared with the ones in the unrestricted case (previous sections), where most of the operators are not strategy-proof. It is not the same story for erosion. We can find profiles that can be manipulated using the erosion strategy (see the running example).

Nevertheless, when d is a distance, Σ is the aggregation function and drastic indexes are considered, $\Delta^{d,\Sigma}$ is strategy-proof if and only if it is erosion strategy-proof:

Proposition 12 *Let d be any distance. If $\Delta^{d,\Sigma}$ is not strategy-proof for i_{d_w} (resp. i_{d_s}), then it is not erosion strategy-proof for index i_{d_w} (resp. i_{d_s}).*

Furthermore, a belief profile E is manipulable by K for i_{d_w} (resp. i_{d_s}) given $\Delta^{d,\Sigma}$ and μ if and only if the manipulation is possible using a complete base $K_\omega \models K$, i.e. there exists $K_\omega \models K$ s.t. $i_{d_w}(K, \Delta_\mu^{d,\Sigma}(E \sqcup \{K_\omega\})) > i_{d_w}(K, \Delta_\mu^{d,\Sigma}(E \sqcup \{K\}))$ (resp. $i_{d_s}(K, \Delta_\mu^{d,\Sigma}(E \sqcup \{K_\omega\})) > i_{d_s}(K, \Delta_\mu^{d,\Sigma}(E \sqcup \{K\}))$).

This result shows that it it enough to focus on each complete base that implies K to determine whether a profile E is manipulable by a base K for i_{d_w}.

Related work

In the propositional merging framework considered in the paper, the beliefs/goals K of each agent induce a two strata partition of the worlds: the models of K are preferred to its countermodels. When agents report full preference relations (that can be encoded in various ways, e.g., explicitly, or by a prioritized belief base, an ordinal conditional function, etc), the aggregation problem consists in defining a global preference relation from individual preference relations. This problem has been addressed for a long time in social choice theory (it can be traced back at least to Condorcet (1785) and Borda (1781)).

In social choice theory (Arrow, Sen, & Suzumura 2002), the strategy-proofness problem has received great attention. One of the more famous result of social choice theory is that there is no strategy-proof preference aggregation procedure. This result is known as Gibbard-Satterthwaite impossibility theorem (Gibbard 1973; Satterthwaite 1975; Moulin 1988).

Since this result has been stated, there has been a lot of work for deriving strategy-proofness results under some restrictions (see (Kelly 1988; Arrow, Sen, & Suzumura 2002) for example). In some sense, our work is relevant to such approaches. Nonetheless, our work is original - as far as we know - from two points of view: on the one hand, the

	#(E)	K	μ	$\Delta^{d_D,f}$	$\Delta^{d_H,\Sigma}$	$\Delta^{d_H,Gmax}$	$\triangle^{C1}$	$\triangle^{C3}$	$\triangle^{C4}$	$\triangle^{C5}$	$\triangle^{\widehat{C1}}$	$\triangle^{\widehat{C3}}$	$\triangle^{\widehat{C4}}$	$\triangle^{\widehat{C5}}$
i_p	$=2$	K_ω	$\top$	√	√	−	−	−	−	−	√	√	√	√
			μ	√	√	−	−	−	−	−	√	−	√	√
		K	$\top$	√	−	−	−	−	−	−	√	√	√	√
			μ	√	−	−	−	−	−	−	√	−	√	√
	>2	K_ω	$\top$	√	√	−	−	−	−	−	−	−	√	−
			μ	√	√	−	−	−	−	−	−	−	√	−
		K	$\top$	√	−	−	−	−	−	−	−	−	√	−
			μ	√	−	−	−	−	−	−	−	−	√	−
i_{d_w}	$=2$	K_ω	$\top$	√	√	−	√	√	−	√	√	√	√	√
			μ	√	√	−	√	−	−	√	√	−	√	√
		K	$\top$	√	√	−	√	√	−	√	√	√	√	√
			μ	√	−	−	√	−	−	−	√	−	√	√
	>2	K_ω	$\top$	√	√	−	√	√	−	√	√	√	√	√
			μ	√	√	−	√	−	−	√	√	−	√	√
		K	$\top$	√	−	−	√	√	−	√	√	√	√	√
			μ	√	−	−	√	−	−	−	√	−	√	−
i_{d_s}	$=2$	K_ω	$\top$	√	√	√	√	√	−	√	√	√	√	√
			μ	√	√	−	√	−	−	−	√	−	√	√
		K	$\top$	√	√	−	√	√	−	√	√	√	√	√
			μ	√	−	−	√	−	−	−	√	−	√	√
	>2	K_ω	$\top$	√	√	−	√	√	−	√	√	√	√	√
			μ	√	√	−	√	−	−	−	√	−	√	−
		K	$\top$	√	−	−	√	√	−	√	√	√	√	√
			μ	√	−	−	√	−	−	−	√	−	√	−

Table 7: The strategy-proofness landscape.

√ means that in the corresponding case (given by the line), the operator (given by the column) is strategy-proof.
− means that in the corresponding case the operator is not strategy-proof.
The first column denotes the satisfaction index under consideration.
The second column states that the results hold for two bases ($=2$) or more (>2).
The third column states that the results hold for complete bases (K_ω) or not (K).
The fourth column states that the results hold when there are integrity constraints (μ) or not ($\top$).

preference relations considered here are two-strata total pre-orders, and not strict total orderings; on the other hand, the result of a merging process is usually not a single world but still a two-strata total pre-order, and the number of models of the merged base is not constrained a priori. That leads to a more complex notion of strategy-proofness. In particular, we must commit to the choice of a satisfaction index, which is not needed when working with preference relations.

A study of strategy-proofness of some merging operators has been carried out in (Meyer, Ghose, & Chopra 2001). The framework considered in this paper is clearly distinct from the one used in our work. Agents may report full preference relations, encoded as stratified belief bases, or equivalently by ordinal conditional functions or κ-functions (one can also look to (Lafage & Lang 2000) for an interesting characterization of weighted preferences aggregation methods). The merging operators under consideration escape Gibbard-Satterthwaite theorem (as well as Arrow theorem) since a commensurability assumption between the agents' preference relations is made (the same remark applies also to possibilistic base merging as defined in (Benferhat *et al.* 2002)). Roughly, commensurability means that we allow to compare the satisfaction degrees of different agents. It means that we do not work with pre-orders, but with a more quantitative framework, where one uses a common (or at least comparable) scale for all agents. It is well-known that the commensurability assumption is sensible in many situations, but when dealing with agents preferences, commensurability must be used carefully (for human agents, it is commonly accepted in social choice theory that this assumption is very strong).

Conclusion and further work

Investigating the strategy-proofness of merging operators is important from a multi-agent perspective whenever some agents can get (part of) the beliefs/goals of the other agents participating to the merging process. When strategy-proofness is not guaranteed, it may be questioned whether the result of the merging process actually represents the beliefs/goals of the group.

In this paper, we have drawn the strategy-proofness landscape for many merging operators, including model-based ones and formula-based ones. While both families are not strategy-proof in the general case, we have shown that several restrictions on the merging framework or on the available strategies may lead to strategy-proofness. As to model-based operators, the choice of a distance appears crucial. Thus, model-based operators are strategy-proof when based on the drastic distance, while they are typically not strategy-proof when based on Dalal distance. Among formula-based merging operators $\triangle^{C1}$ achieves the highest degree of strategy-proofness in the sense that it is strategy-proof for the drastic indexes. Results are summarized in Table 7.

In light of our study, strategy-proofness appears as a property independent from rationality, at least when rationality is captured by the postulates given in (Konieczny & Pino Pérez 1998; 1999). It means that satisfying those rationality postulates neither prevents from manipulability nor implies it. Nevertheless, we can note that arbitration operators (Konieczny & Pino Pérez 2002a), like $\Delta^{d,GMax}$, are more sensitive to manipulation than majority operators, like $\Delta^{d,\Sigma}$. This is easily explained by the fact that arbitration operators are egalitarist: they aim at giving a result that is close to each base of the profile. So a change in a single base can have a real impact on the whole result. Contrastingly, majority operators, that listen to majority wishes to define the resulting base, often do not take into account bases that are not in the majority (this is sometimes called the "majority dictatorship"), hence it is more likely that a change in a single base has no impact on the result of the merging. Strategy-proofness also appears as independent from the computational complexity of query answering from a merged base (see (Konieczny, Lang, & Marquis 2002; 2004)). Hence, strategy-proofness is actually a further dimension that can be used to evaluate and compare merging operators.

This work calls for several perspectives. One of them consists in defining other non-drastic satisfaction indexes. In particular in cases the agent knows that the result of the merging process could not fit her beliefs/goals (for example if her beliefs/goals are not consistent with the integrity constraints), she still can be interested in achieving a result that is as close as possible to her beliefs/goals. Closeness can be captured by a notion of distance, and a possible satisfaction index would be "Dalal index", which could be defined as follows (by homogeneity with the other indexes): $i_{Dalal}(K, K_\Delta) = 1 - \frac{d_H([K],[K_\Delta])}{\#(\mathcal{P})}$.

Another interesting issue is to study the strategy-proofness problem when coalitions are allowed. The question is to know if a group of agents can coordinate for achieving a better result of the merging for all of them. This interesting issue requires more hypotheses on the agents abilities, since it requires communication abilities, in order to allow agents to propose to others to form a coalition, and to coordinate on the base each member of the coalition must give for achieving the wanted result. For this work, it seems that games in coalitional form, studied in game theory (Weber 1994; Greenberg 1994), can provide some interesting notions and results.

A third perspective is to identify the complexity of determining whether a profile can be manipulated by a base given an operator. Indeed, using a merging operator that is not strategy-proof is not necessarily harmful if finding out a strategy is hard. Such a complexity issue has been investigated for voting schemes (Conitzer & Sandholm 2003; Conitzer, Lang, & Sandholm 2003; Conitzer & Sandholm 2002a; 2002b) when individual preferences are given explicitly (which is not the case in our framework). A first result follows easily from Proposition 12: if the distance d between interpretations can be computed in polynomial time in the input size (which is not a strong assumption), determining whether a given profile can be manipulated by a base given $\Delta_\mu^{d,\Sigma}$ and μ is in Σ_2^p.

Acknowledgements

We would like to thank the anonymous reviewers for many helpful comments. The authors have been supported by the

IUT de Lens, the Université d'Artois, the Région Nord/Pas-de-Calais under the TACT-TIC project, and by the European Community FEDER Program.

References

Arrow, K.; Sen, A. K.; and Suzumura, K., eds. 2002. *Handbook of Social Choice and Welfare*, volume 1. North-Holland.

Baral, C.; Kraus, S.; Minker, J.; and Subrahmanian, V. S. 1992. Combining knowledge bases consisting of first-order theories. *Computational Intelligence* 8(1):45–71.

Baral, C.; Kraus, S.; and Minker, J. 1991. Combining multiple knowledge bases. *IEEE Transactions on Knowledge and Data Engineering* 3(2):208–220.

Benferhat, S.; Dubois, D.; Kaci, S.; and Prade, H. 2002. Possibilistic merging and distance-based fusion of propositional information. *Annals of Mathematics and Artificial Intelligence* 34(1-3):217–252.

Borda, J. C. 1781. Mémoire sur les élections au srutin. Histoire de l'Académie Royale des Sciences.

Conitzer, V., and Sandholm, T. 2002a. Complexity of manipulating elections with few candidates. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI'02)*, 314–319.

Conitzer, V., and Sandholm, T. 2002b. Vote elicitation: Complexity and strategy-proofness. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI'02)*, 392–397.

Conitzer, V., and Sandholm, T. 2003. Universal voting protocol tweaks to make manipulation hard. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI'03)*, 781–788.

Conitzer, V.; Lang, J.; and Sandholm, T. 2003. How many candidates are needed to make elections hard to manipulate? In *Proceedings of the Ninth Conference on Theoretical Aspects of Rationality and Knowledge (TARK'03)*, 201–214.

Dalal, M. 1988. Investigations into a theory of knowledge base revision: preliminary report. In *Proceedings of the Seventh American National Conference on Artificial Intelligence (AAAI'88)*, 475–479.

Gibbard, A. 1973. Manipulation of voting schemes. *Econometrica* 41:587–602.

Greenberg, J. 1994. *Handbook of Game Theory*, volume 2. Elsevier. chapter Coalition Structures.

Kelly, J. S. 1988. *Social Choice Theory : An Introduction*. Springer-Verlag.

Konieczny, S., and Pino Pérez, R. 1998. On the logic of merging. In *Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning (KR'98)*, 488–498.

Konieczny, S., and Pino Pérez, R. 1999. Merging with integrity constraints. In *Proceedings of the Fifth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'99)*, LNAI 1638, 233–244.

Konieczny, S., and Pino Pérez, R. 2002a. Merging information under constraints: a qualitative framework. *Journal of Logic and Computation* 12(5):773–808.

Konieczny, S., and Pino Pérez, R. 2002b. On the frontier between arbitration and majority. In *Proceedings of the Eighth International Conference on Principles of Knowledge Representation and Reasoning (KR'02)*, 109–118.

Konieczny, S.; Lang, J.; and Marquis, P. 2002. Distance-based merging: a general framework and some complexity results. In *Proceedings of the Eighth International Conference on Principles of Knowledge Representation and Reasoning (KR'02)*, 97–108.

Konieczny, S.; Lang, J.; and Marquis, P. 2004. DA^2 merging operators. Artificial Intelligence. To appear.

Konieczny, S. 2000. On the difference between merging knowledge bases and combining them. In *Proceedings of the Seventh International Conference on Principles of Knowledge Representation and Reasoning (KR'00)*, 135–144.

Kuipers, T., ed. 1987. *What is closer-to-the-truth?* Rodopi.

Lafage, C., and Lang, J. 2000. Logical representation of preferences for group decision theory. In *Proceedings of the Seventh International Conference on Principles of Knowledge Representation and Reasoning (KR'00)*, 457–468.

Liberatore, P., and Schaerf, M. 1998. Arbitration (or how to merge knowledge bases). *IEEE Transactions on Knowledge and Data Engineering* 10(1):76–90.

Lin, J., and Mendelzon, A. O. 1999. Knowledge base merging by majority. In *Dynamic Worlds: From the Frame Problem to Knowledge Management*. Kluwer.

Marquis de Condorcet. 1785. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Paris.

Meyer, T.; Ghose, A.; and Chopra, S. 2001. Social choice, merging and elections. In *Proceedings of the Sixth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'01)*, 466–477.

Moulin, H. 1988. *Axioms of cooperative decision making*. Econometric society monographs. Cambridge University Press.

Rescher, N., and Manor, R. 1970. On inference from inconsistent premises. *Theory and Decision* 1:179–219.

Revesz, P. Z. 1997. On the semantics of arbitration. *International Journal of Algebra and Computation* 7(2):133–160.

Satterthwaite, M. 1975. Strategy-proofness and Arrow's conditions. *Journal of Economic Theory* 10:187–217.

Weber, R. J. 1994. *Handbook of Game Theory*, volume 2. Elsevier. chapter Games in Coalitional Form.