ECAI 2016
G.A. Kaminka et al. (Eds.)

doi:10.3233/978-1-61499-672-9-372

Beyond IC Postulates: Classification Criteria for Merging Operators

Adrian Haret[1] and **Andreas Pfandler**[1,2] and **Stefan Woltran**[1]

Abstract. Merging is one of the central operations in the field of belief change, which is concerned with aggregating the opinions of individuals. Representation theorems provide a family of merging operators satisfying some natural desiderata for merging beliefs. However, little is known about how these operators can be further distinguished. In the field of social choice, on the other hand, numerous properties have been proposed in order to classify voting rules. In this work, we adapt these properties to the context of merging and investigate how they relate to the standard postulates. Our results thus lead to a more fine-grained classification of merging operators and shed light on the question of which particular merging operator is best suited in a concrete application domain.

1 Introduction

Belief merging studies methods for aggregating the opinions of individuals into a theory which captures the consensus of the agents involved. The standard approach in the literature focuses on the design of merging operators satisfying a set of normative properties. Consensus is then obtained as the theory which comes as close as possible to the agents' expressed beliefs, subject to the limitations expressed by the normative properties [13, 14]. In the field of belief merging the variety of measures of closeness used gives rise via representation theorems to a variety of merging operators with desirable properties. Belief merging differs from voting, as analyzed in (computational) social choice theory (for an overview see, e.g., [1, 16]), in that it does not require the agents to provide full rankings of the alternatives, but only to encode their first choices as logical theories. However, belief merging and voting still share a common goal and methodology, and it is natural to conclude that the two fields can be usefully brought to bear on each other.

One direction of research views voting as a merging task [5, 10], an approach which fits into the larger program of finding suitable logics in which to represent preferences and embed aggregation problems stemming from (computational) social choice [2, 6]. A different approach, which we follow here, looks at merging from a voting perspective and uses the rich set of criteria developed to analyze voting rules in order to classify existing merging operators. Surprisingly, this line of research has received little attention so far. Apart from some interest in strategy-proofness and connections with Arrow's theorem [3, 7, 12, 15], the only other social choice properties that have made their way in the literature on merging are the egalitarian properties discussed in [8]. Notwithstanding, the social choice literature on voting features many other properties whose ideas are relevant in the context of merging, but which have hitherto been left un-addressed. We aim to fill this gap and investigate ways of looking at merging operators that improve upon the fundamental classification into majority and arbitration operators.

Our contribution lies, first of all, in proposing fourteen new properties for merging operators, obtained, mostly, by translating existing properties from the voting literature. In doing so, we contribute to a deeper understanding of merging as a social process, by exploiting a natural analogy between voting and belief merging. Thus, theories to be merged are voters, the merging operator is the voting rule, and interpretations of the propositional variables are the candidates;[3] we also allow for the possibility of a constraint, which limits the range of possible results. In keeping with the merging literature, we take merging operators to be characterized by a core set of properties, known as the IC postulates [13, 14]. Our new properties are meant to extend this characterization by offering more fine-grained criteria for evaluating merging operators. We group the properties according to their character, and offer discussions on the behavior they are intended to model. Second, in the case of each new property, we study its relationship with the core set of IC postulates. When a property is not guaranteed by the IC postulates, we investigate which of the standard operators satisfy the property, give relevant counter-examples, and provide model-based representation results for the most prominent of these properties.

The motivation for proposing new properties is the same as the motivation behind the original IC postulates: we are interested in merging operators that are syntax independent, fair and that respond in expected ways to changes in the input, and we want general principles that capture these properties. Our claim, backed up by the voting literature, is that there are many ways of making these intuitions precise, some of which go beyond the core set of IC postulates.

2 Background

Propositional logic. We work with the language $\mathcal{L}$ of propositional logic over a fixed *alphabet* $\mathcal{P} = \{p_1, \ldots, p_n\}$ of propositional atoms. An interpretation is a set $w \subseteq \mathcal{P}$ of atoms, with the intended meaning that atom p is contained in w if the truth value of p is set to true. The set of all interpretations over $\mathcal{P}$ is denoted by $\mathcal{W}$. We will often represent an interpretation by its corresponding bit-vector of length $|\mathcal{P}|$ (e.g., 101 is the interpretation $\{p_1, p_3\}$). If interpretation w satisfies formula φ, we call w *a model of* φ. We denote the set of models of φ by $[\varphi]$. A pre-order $\leq$ on $\mathcal{W}$ is a binary relation on $\mathcal{W}$ which is reflexive and transitive. We denote by $w_1 < w_2$ the strict part of $\leq$, i.e., $w_1 \leq w_2$

[1] Institute of Information Systems, TU Wien, Austria
[2] School of Economic Disciplines, University of Siegen, Germany

[3] We make an exception to this and treat propositional atoms as candidates when interpretations cannot be reliably seen to fulfil this role. Though this introduces an ambiguity in our notion of "candidate", we view it as a useful step to take in order to capture more voting properties than would otherwise be possible.

but $w_2 \not\leq w_1$. We write $w_1 \approx w_2$ to abbreviate $w_1 \leq w_2$ and $w_2 \leq w_1$. If $\mathcal{M}$ is a set of interpretations, then the set of *minimal elements of $\mathcal{M}$ with respect to* $\leq$ is defined as $\min_{\leq} \mathcal{M} = \{w_1 \in \mathcal{M} \mid \nexists w_2 \in \mathcal{M} \text{ s.t. } w_2 \leq w_1, w_1 \not\leq w_2\}$. By a renaming ρ of the variables we understand a permutation of their names. A renaming ρ applied to any formula, knowledge base or profile changes the propositional variables in it according to ρ. For instance, if ρ swaps only variables p_1 and p_2 among them and $\varphi = p_1 \wedge \neg p_2 \wedge p_3$, then $\rho(\varphi) = p_2 \wedge \neg p_1 \wedge p_3$. We also extend here the notion of renaming to apply to interpretations. Thus, if ρ swaps p_1 and p_2 between them in a formula, then ρ applied to an interpretation swaps the first and second bits in the bit-vector representation. For instance, $\rho(101) = 011$. A transposition τ is a renaming that swaps exactly two elements among each other.

Belief Merging. A *knowledge base* is a finite set of propositional formulas over $\mathcal{L}$. A *profile* is a non-empty finite tuple $E = \langle K_1, \ldots, K_n \rangle$ of consistent, but not necessarily mutually consistent knowledge bases. We denote by $\mathcal{E}$ (resp. $\mathcal{K}$) the set of all profiles (resp. knowledge bases) over $\mathcal{L}$. If E_1 and E_2 are profiles, then $E_1 \sqcup E_2$ is the concatenation of E_1 and E_2. Interpretation w is a model of $K \in \mathcal{K}$ (resp. $E \in \mathcal{E}$) if it is a model of every element in K (resp. E). We denote by $[K]$ and $[E]$ the set of models of K and E, respectively. We write $\bigwedge E$ for $\bigwedge_{K \in E} \bigwedge_{\varphi \in K} \varphi$, $\neg K$ for $\neg \bigwedge K$, and $\neg E$ for $\langle \neg K_1, \ldots, \neg K_n \rangle$. Profiles E_1 and E_2 are *equivalent*, written $E_1 \equiv E_2$, if there exists a bijection $f \colon E_1 \to E_2$ such that for any $K \in E_1$ we have $[K] = [f(K)]$.

A merging operator is a function $\Delta \colon \mathcal{E} \times \mathcal{L} \to \mathcal{K}$, and we write $\Delta_\mu(E)$ instead of $\Delta(E, \mu)$. The formula μ is called *the constraint* and it encodes an external condition which needs to hold in the final result regardless of the input knowledge bases. It can be thought of as a set of legal requirements or limits of feasibility restricting the outcomes of the merging process. Next, logical postulates set out properties which any merging operator Δ should satisfy. An operator satisfying the following postulates is called an IC merging operator [13, 14]:

(IC_0) $\Delta_\mu(E) \models \mu$
(IC_1) If μ is consistent, then $\Delta_\mu(E)$ is consistent
(IC_2) If $\bigwedge E$ is consistent with μ, then $\Delta_\mu(E) \equiv \bigwedge E \wedge \mu$
(IC_3) If $E_1 \equiv E_2$ and $\mu_1 \equiv \mu_2$, then $\Delta_{\mu_1}(E_1) \equiv \Delta_{\mu_2}(E_2)$
(IC_4) If $K_1 \models \mu$ and $K_2 \models \mu$, then $\Delta_\mu(\langle K_1, K_2 \rangle) \wedge K_1$ is consistent iff $\Delta_\mu(\langle K_1, K_2 \rangle) \wedge K_2$ is consistent
(IC_5) $\Delta_\mu(E_1) \wedge \Delta_\mu(E_2) \models \Delta_\mu(E_1 \sqcup E_2)$
(IC_6) If $\Delta_\mu(E_1) \wedge \Delta_\mu(E_2)$ is consistent, then $\Delta_\mu(E_1 \sqcup E_2) \models \Delta_\mu(E_1) \wedge \Delta_\mu(E_2)$
(IC_7) $\Delta_{\mu_1}(E) \wedge \mu_2 \models \Delta_{\mu_1 \wedge \mu_2}(E)$
(IC_8) If $\Delta_{\mu_1}(E) \wedge \mu_2$ is consistent, then $\Delta_{\mu_1 \wedge \mu_2}(E) \models \Delta_{\mu_1}(E) \wedge \mu_2$

Though these postulates lay out what properties $\Delta_\mu(E)$ should have, they do not spell out how to actually construct $\Delta_\mu(E)$, given E and μ. To this end it is useful to focus on so-called *assignments* that map any $E \in \mathcal{E}$ to a pre-order $\leq_E$ on $\mathcal{W}$. We say that such an assignment *represents* a merging operator Δ if $[\Delta_\mu(E)] = \min_{\leq_E}[\mu]$, for any $E \in \mathcal{E}$ and $\mu \in \mathcal{L}$. Konieczny and Pino Pérez [13] have defined the central notion of *syncretic* assignments.

Definition 1. A *syncretic assignment* is a function mapping every $E \in \mathcal{E}$ to a total pre-order $\leq_E$ on $\mathcal{W}$ such that, for any $E, E_1, E_2 \in \mathcal{E}$, $K_1, K_2 \in \mathcal{K}$ and $w_1, w_2 \in \mathcal{W}$ the following conditions hold:

(s_1) If $w_1 \in [E]$ and $w_2 \in [E]$, then $w_1 \approx_E w_2$.
(s_2) If $w_1 \in [E]$ and $w_2 \notin [E]$, then $w_1 <_E w_2$.
(s_3) If $E_1 \equiv E_2$, then $\leq_{E_1} = \leq_{E_2}$.
(s_4) If $w_1 \in [K_1]$, then there is $w_2 \in [K_2]$ s.t. $w_2 \leq_{\{K_1, K_2\}} w_1$.
(s_5) If $w_1 \leq_{E_1} w_2$ and $w_1 \leq_{E_2} w_2$, then $w_1 \leq_{E_1 \sqcup E_2} w_2$.
(s_6) If $w_1 \leq_{E_1} w_2$ and $w_1 <_{E_2} w_2$, then $w_1 <_{E_1 \sqcup E_2} w_2$.

The classical result below characterizes all IC merging operators in terms of syncretic assignments.

Theorem 1. *A merging operator Δ is an IC merging operator iff there is a syncretic assignment which represents it.*

Specifying concrete merging operators is usually done via a notion of distance (that induces pre-orders $\leq_{K_i}$) and an aggregation function (which combines the individual rankings $\leq_{K_i}$ into a final pre-order $\leq_E$). More precisely: a *pseudo-distance* is a function $d \colon \mathcal{W} \times \mathcal{W} \to \mathbb{R}_+$ such that, for any $w_1, w_2 \in \mathcal{W}$, (i) $d(w_1, w_2) = d(w_2, w_1)$ and (ii) $d(w_1, w_2) = 0$ if and only if $w_1 = w_2$. An *aggregation function* is a function f such that, for any $x_1, \ldots, x_n, x, y \in \mathbb{R}_+$ and any permutation π, (i) if $x \leq y$, then $f(x_1, \ldots, x, \ldots, x_n) \leq f(x_1, \ldots, y, \ldots, x_n)$, (ii) $f(x_1, \ldots, x_n) = 0$ if and only if $x_1 = \cdots = x_n = 0$, (iii) $f(x) = x$, and (iv) $f(x_1, \ldots, x_n) = f(\pi(x_1), \ldots, \pi(x_n))$.

The *Hamming distance* between interpretations w and w' is defined as $d_H(w, w') = |(w \setminus w') \cup (w' \setminus w)|$; the *drastic distance* between w and w' is given as $d_D(w, w') = 0$ if $w = w'$ and 1 otherwise. The minimal distance between interpretations and models of K_i yields $\leq_{K_i}$. The distance $d(w, K_i)$ between w and K_i is computed by taking the minimal distance between w and all $w' \in [K_i]$. Now, the pre-order $\leq_{K_i}$ is defined by saying that $w \leq_{K_i} w'$ if $d(w, K_i) \leq d(w', K_i)$. For aggregating the rankings, common functions are summation (Σ), $GMAX$ and $GMIN$. For d_H this gives us the operators $\Delta^{d_H, \Sigma}$, $\Delta^{d_H, GMAX}$ and $\Delta^{d_H, GMIN}$. These operators are all distinct, in the sense that they may give different results on the same input. On the other hand, operators defined using the drastic distance are all equivalent, in the sense that $\Delta_\mu^{d_D, \Sigma}(E) \equiv \Delta_\mu^{d_D, GMAX}(E) \equiv \Delta_\mu^{d_D, GMIN}(E)$, for any $E \in \mathcal{E}$ and $\mu \in \mathcal{L}$. We thus denote these operators by Δ^{d_D}. In general, we will write Δ^{d_H} and Δ^{d_D} when our results hold with all of the three aggregation functions presented. For details, see [13, 14].

Example 1. Consider three reviewers who are part of a conference committee. They have to arrive at a decision concerning three papers they have been assigned, in a process that requires combining their individual (perhaps mutually inconsistent) beliefs about which of the papers should be accepted or rejected. The acceptance of each paper is represented by a propositional atom: p_i means that paper i is accepted, for $i \in \{1, 2, 3\}$. The opinions of the three reviewers are encoded by three knowledge bases, as follows: $K_1 = \{p_1 \wedge p_2 \wedge \neg p_3\}$, $K_2 = \{\neg p_1 \wedge \neg p_2\}$, $K_3 = \{p_1 \wedge p_3\}$. In other words, Reviewer 1 thinks only Papers 1 and 2 should be accepted, Reviewer 2 thinks Papers 1 and 2 should be rejected but has no stated opinion on Paper 3, and Reviewer 3 thinks Papers 1 and 3 should be accepted. Additionally, the rule for their committee is that not all papers can be accepted. This rule can be encoded by the constraint $\mu = \neg(p_1 \wedge p_2 \wedge p_3)$.

Thus, if $E = \langle K_1, K_2, K_3 \rangle$ is the profile, the task is to compute $\Delta_\mu(E)$. We illustrate the operators $\Delta^{d_H, \Sigma}$, $\Delta^{d_H, GMAX}$, $\Delta^{d_H, GMIN}$ discussed above. First we compute a pre-order $\leq_{K_i}$ on $\mathcal{W}$ for each K_i based on the distance $d(w, K_i)$. In our example (using Hamming distance d_H), we obtain $d(010, K_3) = \min\{d_H(010, w') \mid w' \in [K_3]\} = \min\{d_H(010,101), d_H(010,111)\} = \min\{3, 2\} = 2$. The complete set of distances is featured in Table 1. The next step is to combine the pre-orders $\leq_{K_i}$ into a new pre-order, reflecting

w	K_1	K_2	K_3	Σ	$GMAX$	$GMIN$
000	2	0	2	4	(2,2,0)	(0,2,2)
001	3	0	1	4	(3,1,0)	(0,1,3)
010	1	1	2	4	(2,1,1)	(1,1,2)
011	2	1	1	4	(2,1,1)	(1,1,2)
100	1	1	1	3	(1,1,1)	(1,1,1)
101	2	1	0	3	(2,1,0)	(0,1,2)
110	0	2	1	3	(2,1,0)	(0,1,2)
111	1	2	0	3	(2,1,0)	(0,1,2)

Table 1: Distances and aggregated values for Example 1

the consensus opinion. We use an aggregation function (in this case Σ, $GMAX$ and $GMIN$) to obtain the final ranking $\leq_E$. The aggregation function Σ adds the numbers interpretation-wise, and the final ranking $\leq_E^\Sigma$ is determined by the order of the final levels for each interpretation. The aggregation functions $GMAX$ and $GMIN$ order the vector of levels for each interpretation in descending and ascending order, respectively. Then we determine $\leq_E^{GMAX}$ and $\leq_E^{GMIN}$ by ordering the vectors lexicographically. Finally, we pick from the models of μ (highlighted in grey in Table 1) the ones with minimal levels in the final ranking: $[\Delta_\mu^{d_H,\Sigma}(E)] = \{100, 101, 110\}$, $[\Delta_\mu^{d_H,GMAX}(E)] = \{100\}$, $[\Delta_\mu^{d_H,GMIN}(E)] = \{101, 110\}$.

We can now interpret this result back in propositional logic. For instance, $\Delta_\mu^{d_H,\Sigma}(E) \equiv \{p_1 \wedge \neg(p_2 \wedge p_3)\}$, thus saying that Paper 1 should be accepted, but Papers 2 and 3 cannot be accepted together. Notice that this result is not resolute, as it does not tell which of Papers 2 or 3 should be accepted, if any. On the other hand, $\Delta_\mu^{d_H,GMAX}(E) \equiv \{p_1 \wedge \neg p_2 \wedge \neg p_3\}$, thus saying that only Paper 1 should be accepted.

In the next section we will revisit this example several times to illustrate and clarify the definitions of the properties.

Voting Theory. Let C be a finite set of *candidates* with $|C| = m$ and $V = \{1, 2, \ldots, n\}$ be a finite set of *voters*. The preference of a voter is modelled as a total order over C, the *vote* $\succ$. The top-ranked candidate of $\succ$ is at position 1, the successor at position 2, ..., and the last-ranked candidate is at position m. A collection of preference relations $\mathcal{P} = \langle \succ_1, \ldots, \succ_n \rangle$ is called a *preference profile*. A voter i *prefers* candidate c over candidate c' if $c \succ_i c'$. An *election* is given by $E = (C, V, \mathcal{P})$. A *voting correspondence* $\mathcal{F}$ is a mapping from an election E to a non-empty subset of the candidates $W \subseteq C$, i.e., the *winners* of the election. We denote a preference profile comprising pre-orders instead of total orders by $\langle \leq_1, \ldots, \leq_n \rangle$. For more details, see [1] and [16, in particular Chapter 4].

3 Properties for Belief Merging

In this section we present a number of natural properties for belief merging, several of which stem from the (computational) social choice literature, where they are typically applied to voting procedures. We group these properties according to common themes of interest in the Knowledge Representation literature.

Syntax Independence

These properties require that the outcome does not depend on how knowledge is encoded. In other words, the concrete syntactic formulation of the profile should not affect the result of the merging process. Note that IC_3 already ensures syntax independence to some degree. However, not all properties defined here are implied by IC_3 and hence are more restrictive.

Anonymity. A voting system satisfies anonymity if the winner cannot be changed by permuting the votes in the profile. In a merging scenario, we denote by $\pi(E) = \langle K_{\pi(1)}, \ldots, K_{\pi(n)} \rangle$ the profile obtained by changing the order of the knowledge bases in E in accordance with a permutation $\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$. Anonymity is then defined as follows:[4]

(Anonymity) $\Delta_\mu(E) \equiv \Delta_\mu(\pi(E))$.

In Example 1, Anonymity requires that $\Delta^{d_H,\Sigma}$, $\Delta^{d_H,GMAX}$ and $\Delta^{d_H,GMIN}$ produce the same result, respectively, when the profile is $\langle K_1, K_2, K_3 \rangle$, or $\langle K_2, K_1, K_3 \rangle$, or is any other permutation of the knowledge bases. One can see that Anonymity is satisfied by the operators mentioned, as the final result does not depend on the order in which the pre-orders are aggregated.

Neutrality. In a voting scenario, neutrality requires that if two candidates are swapped in all votes, then they are also swapped in the result. The purpose is to ensure that all candidates are treated equally in the determination of the winners, i.e., their name does not matter. In a merging scenario, we have to enforce that renaming variables does not affect the merging outcome. We define neutrality as follows:

(Neutrality) $\rho(\Delta_\mu(E)) \equiv \Delta_{\rho(\mu)}(\rho(E))$.

Under the renaming ρ that swaps p_1 and p_2 among them, the knowledge bases and constraint from Example 1 become: $\rho(K_1) = \{p_2 \wedge p_1 \wedge \neg p_3\}$, $\rho(K_2) = \{\neg p_2 \wedge \neg p_1\}$, $\rho(K_3) = \{p_2 \wedge p_3\}$ and $\rho(\mu) = \neg(p_2 \wedge p_1 \wedge p_3)$. Computing $\Delta_{\rho(\mu)}^{d_H,\Sigma}(\rho(E))$ we get that $\Delta_{\rho(\mu)}^{d_H,\Sigma}(\rho(E)) \equiv \{p_2 \wedge \neg(p_1 \wedge p_3)\} \equiv \rho(\Delta_\mu^{d_H,\Sigma}(E))$.

Entity resolution. Suppose that, at some point in the knowledge modelling process, different variables, e.g., p and q, are discovered to encode the same concept. The knowledge engineer would want to incorporate this equivalence in the merging outcome. One way to do this is going through the knowledge bases and the constraint and renaming p to q. This is a laborious and invasive operation, which might be infeasible if access to the knowledge bases is limited or if the knowledge bases are provided by the agents just in time before the merging process. Another way is to add the equivalence $p \leftrightarrow q$ directly to μ. The property we propose explores the relationship between these two operations and requires that all solutions of the latter operation are also solutions of the former. We denote by $\mu^{p/q}$ and $K^{p/q}$ the formula and knowledge base obtained from μ and K, respectively, by replacing every occurrence of p with q. We denote by $E^{p/q}$ the profile obtained from E by replacing every knowledge base K in E by $K^{p/q}$, if $K^{p/q}$ is consistent; if $K^{p/q}$ is inconsistent, we remove it.

(Entity resolution) $\Delta_{\mu \wedge (p \leftrightarrow q)}(E) \models \Delta_{\mu^{p/q}}(E^{p/q})$.

Entity resolution has no direct equivalent in the voting scenario. Nonetheless, we believe it is worth investigating, as it bears some resemblance to Independence of clones (see below) and is motivated by a similar intuition: alternatives that are in some sense redundant should not skew the vote in their favour. In Example 1, we obtain that $\Delta_{\mu \wedge (p_1 \leftrightarrow p_2)}^{d_H}(E) \equiv \{p_1 \wedge p_2 \wedge \neg p_3\}$ (regardless of the aggregation function used). Replacing every occurrence of p_1 with p_2 in μ and E leaves us with $K_1^{p_1/p_2} = \{p_2 \wedge \neg p_3\}$, $K_2^{p_1/p_2} = \{\neg p_2\}$, $K_3^{p_1/p_2} = \{p_2 \wedge p_3\}$ and $\mu^{p_1/p_2} = \neg(p_2 \wedge p_3)$, and $\Delta_{\mu^{p_1/p_2}}^{d_H}(E^{p_1/p_2}) \equiv \{p_2 \wedge \neg p_3\}$. Clearly, in this case we have that $\Delta_{\mu \wedge (p_1 \leftrightarrow p_2)}^{d_H}(E) \models \Delta_{\mu^{p_1/p_2}}^{d_H}(E^{p_1/p_2})$. However, we show in Section 4 that this does not hold in general.

[4] Here and in the following, variables such as E, K and μ are understood to be universally quantified, unless explicitly mentioned otherwise.

Fairness

The second group proposes a set of *fairness* properties, stemming from the intuition that all knowledge bases and all variables should be treated "equally" in the merging process. Fairness is featured in the IC postulates (through IC_4), but our proposals show that there is a wider range of constraints to consider.

Non-dictatorship. In a voting scenario, this property is satisfied if there is no single voter that alone determines the outcome of the election, and is usually featured as a key requirement that any reasonable voting method should satisfy. Here we distinguish between two notions: $\mathsf{Non\text{-}Dictatorship}_1$ is in the spirit of the property usually found for voting, while $\mathsf{Non\text{-}Dictatorship}_2$ has a more semantic flavour.

($\mathsf{Non\text{-}Dictatorship}_1$) There is no integer i such that for any $E \in \mathcal{E}$ and $\mu \in \mathcal{L}$, it holds that $\Delta_\mu(E) \equiv \Delta_\mu(\langle K_i \rangle)$.

($\mathsf{Non\text{-}Dictatorship}_2$) There is no $K \in \mathcal{K}$ such that for any $E \in \mathcal{E}$ and $\mu \in \mathcal{L}$, it holds that if a $K' \in \mathcal{K}$ occurs in E with $K' \equiv K$, then $\Delta_\mu(E) \equiv \Delta_\mu(\langle K' \rangle)$, for any $\mu \in \mathcal{L}$.

Property $\mathsf{Non\text{-}Dictatorship}_2$ specifies that there is no knowledge base in the semantic sense (i.e., a specific set of beliefs modulo logical equivalence) which, if present in a profile, unilaterally determines the merging outcome. In a voting setting, this is equivalent to saying there is no *ranking of alternatives* (think of it as a magic key) which, if submitted by some voter, decides the winners. Non-dictatorship has been mentioned before in relation to Arrow's theorem [12], though it has not been formalized explicitly.

Pareto consistency. A voting system is Pareto-consistent if whenever all voters prefer a candidate c_i over some candidate c_j, then c_i is preferred over c_j in the result. A stronger version stipulates that no candidates other than those preferred by all voters should appear in the result. In a merging scenario, we correspondingly distinguish between a weak and a strong version of Pareto consistency.

($\mathsf{Weak\ Pareto}$) $\Delta_\mu(\langle K_1 \rangle) \wedge \cdots \wedge \Delta_\mu(\langle K_n \rangle) \models \Delta_\mu(\langle K_1, \ldots, K_n \rangle)$.

($\mathsf{Strong\ Pareto}$) If $\Delta_\mu(\langle K_1 \rangle) \wedge \cdots \wedge \Delta_\mu(\langle K_n \rangle)$ is consistent, then $\Delta_\mu(\langle K_1, \ldots, K_n \rangle) \models \Delta_\mu(\langle K_1 \rangle) \wedge \cdots \wedge \Delta_\mu(\langle K_n \rangle)$.

Replacing IC_5 and IC_6 in the IC postulates with Weak Pareto and Strong Pareto yields what is called a *pre*-IC *merging operator*. In [8] it has already been noted that any IC merging operator is also a pre-IC merging operator. Pareto conditions also occur in [4] in connection to a related set of operators called *fusion operators*.

Citizen's sovereignty. In a voting scenario, citizen's sovereignty requires that for any candidate c there is at least one election such that c is the winner. In other words, no candidate is disadvantaged by the voting system *per se*. In a merging scenario, we require that no formula is disadvantaged by the operator *per se*.

($\mathsf{Citizen's\ sovereignty}$) For any formula φ there exist $E \in \mathcal{E}$ and $\mu \in \mathcal{L}$ such that $\Delta_\mu(E) \equiv \varphi$.

SC-Majority. This property requires that a candidate c is a winner whenever more than half of the voters have c as their most preferred candidate. Considering a formula φ as a set of candidates (i.e., φ's models) and the knowledge bases K_i as the voters, we have:

($\mathsf{SC\text{-}Majority}$) If $\varphi \in \mathcal{L}$ is consistent and $\varphi \models \Delta_\mu(\langle K_i \rangle)$ for a majority of $i \in \{1, \ldots, n\}$, then $\varphi \models \Delta_\mu(E)$.

In Example 1 there is no consistent formula φ such that $\varphi \models \Delta_\mu(\langle K_i \rangle)$ for a majority of $i \in \{1, \ldots, n\}$. Hence, when we view $\Delta_\mu^{d_H}(E)$ as an election over the models of μ, there is no majority winner. However, merging the same profile under the constraint $\mu' = (p_1 \oplus p_2) \wedge \neg p_3$, we observe that μ' is a majority winner but $\Delta_{\mu'}^{d_H}(E) \equiv \{p_1 \wedge \neg p_2 \wedge \neg p_3\}$. Clearly, though, $\mu' \not\models \Delta_{\mu'}^{d_H}(E)$.

Condorcet criterion. In a voting scenario the Condorcet criterion is satisfied if the voting system selects the Condorcet winner, if it exists. The Condorcet winner is a candidate that beats every other candidate in pairwise majority comparisons. In a merging scenario, our proposal is to define majority comparisons in terms of complete formulas.[5] We opted to present this version here as it directly captures the intuition of the Condorcet winner from the voting scenario. In a more extensive treatment of the topic we would present it alongside an equivalent simpler version.

Definition 2. Given a merging operator Δ, $E \in \mathcal{E}$, $\mu \in \mathcal{L}$ and two complete formulas $\varphi_1, \varphi_2 \in \mathcal{L}$ such that $\varphi_1 \models \mu$ and $\varphi_2 \models \mu$, a *head-to-head election between* φ_1 *and* φ_2 occurs as follows: for every K_i in E, we say that φ_1 wins over φ_2 with respect to K_i if $\Delta_{\varphi_1 \vee \varphi_2}(\langle K_i \rangle) \wedge \varphi_1$ is consistent and $\Delta_{\varphi_1 \vee \varphi_2}(\langle K_i \rangle) \wedge \varphi_2$ is inconsistent. If both $\Delta_{\varphi_1 \vee \varphi_2}(\langle K_i \rangle) \wedge \varphi_1$ and $\Delta_{\varphi_1 \vee \varphi_2}(\langle K_i \rangle) \wedge \varphi_2$ are consistent, we say that φ_1 and φ_2 are tied with respect to K_i. We denote by $W_E(\varphi_1, \varphi_2)$ the number of wins of φ_1 over φ_2 in E. Finally, we say that φ_1 wins over φ_2 in a head-to-head election over E if $W_E(\varphi_1, \varphi_2) \geq W_E(\varphi_2, \varphi_1)$.[6] A complete formula φ such that $\varphi \models \mu$ is a *weak Condorcet winner with respect to* E *and* μ if for any complete formula $\varphi' \models \mu$ such that $\varphi \not\equiv \varphi'$, it holds that φ wins over φ' in a head-to-head election over E.

($\mathsf{Condorcet's\ criterion}$) If φ is a weak Condorcet winner with respect to E and μ, then $\varphi \models \Delta_\mu(E)$.

According to our definition, a weak Condorcet winner on formulas can be shown to coincide with the more familiar notion of a weak Condorcet winner from voting theory, by viewing the set of pre-orders $\langle \leq_{\langle K_1 \rangle}, \ldots, \leq_{\langle K_n \rangle} \rangle$ in a syncretic assignment as a voting profile where $[\mu]$ is the set of candidates (see Theorem 3 and the $\mathsf{s_{Con}}$ property). Applying this result here, we consider merging the profile E from Example 1 under a constraint μ' such that $[\mu'] = \{000, 001, 010, 100\}$ and using Hamming distance and Σ as aggregation function. We obtain the same table of distances from Example 1, except that we restrict our attention to the models of μ'. Table 2 records the number of wins of each interpretation in $[\mu']$ over the other in the resulting voting profile: an entry of k in row i and column j means that interpretation w_i has k wins over interpretation w_j. For instance, 000 has only one win over 001 (namely, $000 <_{K_1} 001$). Likewise, 001 has only one win over 000 (namely, $001 <_{K_3} 000$).[7] Obviously, from a voting perspective it does not make sense to compare an interpretation to itself, thus the entries on the diagonal are marked with "-". Inspection of Table 2 then shows that 001 and 100 are the only models that do not lose to any other interpretation, which means that they are the weak Condorcet winners in this profile. Hence the formulas $\varphi_1 = \neg p_1 \wedge \neg p_2 \wedge p_3$ and $\varphi_2 = p_1 \wedge \neg p_2 \wedge \neg p_3$ are the corresponding weak Condorcet winners.

[5] Complete formulas have exactly one model.

[6] We opted to go with the weak form of Condorcet winner because we did not wish to restrict the set of winners to have exactly one model. However, we could define a strong notion of Condorcet winner by requiring the inequality between $W_E(\varphi_1, \varphi_2)$ and $W_E(\varphi_2, \varphi_1)$ to be strict and our analysis would still go through. For reasons of space we omit this here.

[7] We do not count the tie $000 \approx_{K_2} 001$.

	000	001	010	100
000	-	1	1	1
001	1	-	2	1
010	1	1	-	0
100	2	1	1	-

Table 2: Computing the Condorcet winners

On the other hand, $\Delta^{d_H,\Sigma}_{\mu'}(E) \equiv \{p_1 \wedge \neg p_2 \wedge \neg p_3\}$ and it is clear that $\varphi_1 \not\models \Delta^{d_H,\Sigma}_{\mu'}(E)$ and thus $\Delta^{d_H,\Sigma}_{\mu'}(E)$ does not select (all) the weak Condorcet winners of this profile. In Section 4 we will show that this observation generalizes to the other merging operators.

Intuitive Response to Profile Change

This group of properties ensures that changes in the knowledge base produce an intuitive change of the outcome. Having an *intuitive response* of the formalism is particularly important for knowledge engineers as it reduces unnatural behavior and makes the effects of changes in the knowledge bases easier to grasp.

Monotonicity. A voting system is monotone if the winner of an election cannot be turned into a non-winner by improving its rank in some of the votes. In the context of merging, we propose:

(Monotonicity) $\Delta_\mu(E_1 \sqcup E_2) \wedge \Delta_\mu(E_3) \models \Delta_\mu(E_1 \sqcup E_3)$.

The intuition behind this formalization stems from seeing the models of $\Delta_\mu(E)$ as the winners in the election where the models of μ are candidates and the knowledge bases in E are the voters. Thus, if any candidates elected by the profile $E_1 \sqcup E_2$ are also elected by the profile E_3 alone, then monotonicity would require that the same candidates should also be elected when we replace E_2 with E_3 in $E_1 \sqcup E_2$. The idea, to put it succinctly, is that a winner stays a winner, if its position is only increased in the votes.

To illustrate the property, consider the knowledge bases in Example 1 and a constraint μ' such that $[\mu'] = \{011, 100\}$. Then $100 \in [\Delta^{d_H}_{\mu'}(\langle K_1, K_2\rangle)]$. In other words, the interpretation 100 is a winner in an election where K_1 and K_2 are the voters and the interpretations 011 and 100 (as models of μ') are the sole candidates. We also see, by consulting Table 1, that $100 \in [\Delta^{d_H}_{\mu'}(\langle K_3\rangle)]$, i.e., the voter K_3 counts 100 among its most preferred states. Monotonicity would then require that replacing K_2 with K_3 in the profile $\langle K_1, K_2\rangle$ would not harm the position of 100 in the result. And indeed, we have that $100 \in [\Delta^{d_H}_{\mu'}(\langle K_1, K_3\rangle)]$, showing that Monotonicity is satisfied in this particular instance. However, in Section 4 we show that Monotonicity is not satisfied in general by Δ^{d_H} and Δ^{d_D}.

Participation. A voting system satisfies participation (also known as the *no-show paradox*) if it is not possible to change the winner from candidate c_i to candidate c_j by adding a vote in which candidate c_i is strictly preferred to candidate c_j. In a merging scenario, we consider adding a knowledge base K to a given profile E and require that $\Delta_\mu(E \sqcup \langle K\rangle)$ should not be 'worse' than $\Delta_\mu(E)$ with respect to K.

(Participation) If $\Delta_\mu(E) \wedge K$ is consistent, then $\Delta_\mu(E) \wedge K \models \Delta_\mu(E \sqcup \langle K\rangle)$.

In Example 1, take $\mu' = \neg p_2 \wedge p_3$, with $[\mu'] = \{001, 101\}$. We have that $[\Delta^{d_H,\Sigma}_{\mu'}(\langle K_1, K_2\rangle)] = \{101\}$. In other words, if Reviewers 1 and 2 decided alone, then 101 would be their most preferred state, as chosen by $\Delta^{d_H,\Sigma}$. Notice that 101 is also a model of K_3, i.e., Reviewer 3 also has 101 among its most preferred states. We can imagine Reviewer 3 has a choice: she can either express her opinions, or stand by as a passive observer. Now, if there was a possibility that weighing in with her true opinions would *decrease* the chance that 101 appears in the result, then Reviewer 3 would have an incentive to keep her opinion to herself. This does not happen, as $101 \in [\Delta^{d_H,\Sigma}_{\mu'}(\langle K_1, K_2, K_3\rangle)]$. Hence it is safe for Reviewer 3 to weigh in on the reviewing process with her true opinions. We would want all merging operators to emulate this property, as it incentivizes agents to participate with their honest opinions.

Reversal symmetry. This property holds in a voting system if the unique winner of an election can be turned into a non-winner by reversing all votes. In a merging scenario, we interpret the condition of having a unique winner as the outcome of merging being a complete formula, and we take reversing the vote to mean that every knowledge base is replaced with its negation, as defined in Section 2. Notice that we require the outcome to be a complete formula to reflect the requirement of a unique winner in the voting setting.

(Reversal symmetry) If $\Delta_\mu(E)$ is a complete formula and μ has more than one model, then $\Delta_\mu(E) \not\models \Delta_\mu(\neg E)$.

In Example 1, replacing every knowledge base with its negation gives $\neg K_1 = \{\neg(p_1 \wedge p_2 \wedge \neg p_3)\}$, $\neg K_2 = \{\neg(\neg p_1 \wedge \neg p_2)\}$ and $\neg K_3 = \{\neg(p_1 \wedge p_3)\}$. Merging these knowledge bases with $\Delta^{d_H,GMAX}$ under the constraint μ (from the example) produces the result $\Delta^{d_H,GMAX}_\mu(\neg E) = \{010, 011, 100, 101\}$. Thus, $\Delta^{d_H,GMAX}_\mu(E) \models \Delta^{d_H,GMAX}_\mu(\neg E)$, and hence $\Delta^{d_H,GMAX}$ does not satisfy Reversal symmetry. In Section 4 it is shown that this result extends to other merging operators as well.

Resolvability. In a voting scenario, resolvability (see, e.g., [17]) requires that any winner can be made the unique winner by adding a single vote. In a merging scenario, we require that we can refine the output of merging as much as we desire by adding just one knowledge base to E.

(Resolvability) For any $\varphi \in \mathcal{L}$ such that $\varphi \models \Delta_\mu(E)$, there is a $K \in \mathcal{K}$ such that $\Delta_\mu(E \sqcup \langle K\rangle) \equiv \varphi$.

It has been pointed out in Section 2 that the output of a merging operator is not always resolute, in the sense of selecting a completely specified state of affairs. In Example 1 we got that $\Delta^{d_H,\Sigma}_\mu(E) \equiv \{p_1 \wedge \neg(p_2 \wedge p_3)\}$, thus saying that Paper 1 should be accepted while Papers 2 and 3 cannot be accepted together, but not giving any additional information on which (if any) of Papers 2 and 3 should be accepted. This is because merging operators are designed to offer a solution based on the available information, and that might be insufficient to decide between a set of alternatives. However, in certain circumstances, such as the one offered in Example 1, we might want an answer that settles the question definitively. In such a case, it is reasonable to do so by eliciting more information from the agents involved. The Resolvability property analyzes the possibility that the result can be refined enough by adding a single vote, so as to settle on a decision regarding every option. In Example 1 we can settle on the decision where, for instance, only Paper 1 is accepted by adding the knowledge base $K_4 = \{p_1 \wedge \neg p_2 \wedge \neg p_3\}$ to the profile. Notice that our definition of resolvability does not require the operator to be resolute.

Independence of clones. In a voting scenario, we say that two candidates are clones if they are ranked next to each other in any vote of the election. A voting system is independent of clones if a non-winning candidate cannot be made a winner by adding clones to the election.

In a merging scenario, as it does not make too much sense to think of introducing new interpretations, we think of clones as new *variables* that are equivalent to existing ones. Thus, given a merging profile $E = \langle K_1, \dots, K_n \rangle$, a propositional variable p and a set of "new" propositional variables $\mathcal{Q} \subset \mathcal{P}$ not appearing in $K_1, \dots, K_n$ (clones of p), we denote by $E^{p,\mathcal{Q}}$ the profile obtained by adding the formula $\bigwedge_{q \in \mathcal{Q}}(p \leftrightarrow q)$ to every knowledge base K_i contained in E. Independence of clones for merging operators is formulated as follows:

(Independence of clones) If every $K \in E^{p,\mathcal{Q}}$ is consistent, then $\Delta_\mu(E) \equiv \Delta_\mu(E^{p,\mathcal{Q}})$.

Consider merging the knowledge bases $\{p\}$ and $\{\neg p\}$ with the operator $\Delta^{d_H,\Sigma}$: we get that $\Delta_\top^{d_H,\Sigma}(\langle\{p\},\{\neg p\}\rangle) \equiv \top$. Adding a clone q of p gives us that $\Delta_\top^{d_H,\Sigma}(\langle\{p, p \leftrightarrow q\},\{\neg p, p \leftrightarrow q\}\rangle) \equiv \top$. Adding a clone to the profile does not change the final result and this seems fitting, as introducing the new information regarding q does not change the agents' beliefs regarding the "main" issue, represented by p. Hence, one would like to see this behaviour reproduced more generally. However, Independence of clones as we have formulated it is a very strong property. Thus, adding a clone p_4 for p_1 in Example 1 produces the result that $\Delta_\mu^{d_H,\Sigma}(E) \equiv \{p_1 \wedge \neg(p_2 \wedge p_3) \wedge p_4\}$, $\Delta_\mu^{d_H,GMAX}(E) \equiv \{p_1 \wedge \neg p_2 \wedge \neg p_3 \wedge p_4\}$, $\Delta_\mu^{d_H,GMIN}(E) \equiv \{p_1 \wedge p_2 \wedge \neg p_3 \wedge p_4\}$. Obviously Independence of clones is not satisfied here, and Section 4 shows that this result generalizes.

Modularity

Modularity properties capture circumstances where a profile can be decomposed into sub-profiles while preserving the merging result.

Consistency. In a voting scenario, consistency requires that if an election E is arbitrarily divided into sub-elections $E_1, \dots, E_n$ and if candidate c is a winner in all of the sub-elections $E_1, \dots, E_n$, then c is also a winner of E. For merging we formulate consistency as follows:

(Consistency) For any partition $E_1, \dots, E_n$ of E it holds that $\Delta_\mu(E_1) \wedge \dots \wedge \Delta_\mu(E_n) \models \Delta_\mu(E)$.

Observe that Consistency and Weak Pareto do not coincide, as Consistency is stronger than Weak Pareto.

Stability

These properties are subtly different to those describing intuitive response to profile change: they model modifications of the knowledge bases which should *not* affect the result of the merging process.

Homogeneity. A voting procedure satisfies homogeneity if for any $k \geq 1$ and any election, the result cannot be changed by "repeating" each vote k times. In a merging scenario we require that the outcome of merging does not change if we expand the profile by adding multiple copies of itself. That is, the absolute "weights" of the knowledge bases are not relevant—rather it is the relative weights that matter.

(Homogeneity) $\Delta_\mu(E) \equiv \Delta_\mu(E \sqcup \dots \sqcup E)$.

Self-agreement. We require that the merging outcome is not disrupted if we add it back to E and merge the new profile.

(Self-agreement) $\Delta_\mu(E \sqcup \langle \Delta_\mu(E) \rangle) \equiv \Delta_\mu(E)$.

4 Relationship with IC postulates

In this section we analyze the properties introduced in Section 3. The results are summarized in Table 3. A significant number of the properties we introduced turn out to follow directly from the IC postulates, whereas there are some that hold for certain operators only.

Property	IC	Δ^{d_H}	Δ^{d_D}
Anonymity	✓	✓	✓
Neutrality		✓	✓
Entity resolution		×	✓
Non-Dictatorship$_1$, Non-Dictatorship$_2$	✓	✓	✓
Weak Pareto*	✓	✓	✓
Strong Pareto*	✓	✓	✓
Citizen's sovereignty	✓	✓	✓
SC-Majority	↯	×	×
Condorcet's criterion		×	✓
Monotonicity		×	×
Participation	✓	✓	✓
Reversal symmetry		×	✓
Resolvability	✓	✓	✓
Independence of clones		×	×
Consistency	✓	✓	✓
Homogeneity	✓	✓	✓
Self-agreement	✓	✓	✓

Table 3: Summary of results. In the IC column, ✓ indicates that the property is implied by the IC postulates, and ↯ indicates that the property is inconsistent with the IC postulates. The last two columns indicate whether the property holds for operators based on Hamming distance (Δ^{d_H}) and drastic distance (Δ^{d_D}). Results for properties marked by $*$ have already been studied [8].

Theorem 2. *Anonymity, Non-Dictatorship$_1$, Non-Dictatorship$_2$, Weak Pareto, Strong Pareto, Citizen's sovereignty, Participation, Resolvability, Consistency, Homogeneity and Self-agreement follow from the IC postulates.*

Proof. For Anonymity take the bijection $f(K_i) = K_{\pi(i)}$ between E and $\pi(E)$ and apply IC_3. Non-Dictatorship$_1$ follows from Anonymity, as in the classical voting scenario. For Non-Dictatorship$_2$, suppose K_1 is a dictator for Δ. Choose a (consistent) K_2 such that $\bigwedge K_1 \wedge \bigwedge K_2$ is inconsistent, and $\mu = \bigwedge K_1 \vee \bigwedge K_2$. Clearly, $\Delta_\mu(\langle K_1, K_2\rangle) \wedge \bigwedge K_1$ is consistent, and thus (by IC_4) it holds that $\Delta_\mu(\langle K_1, K_2\rangle) \wedge \bigwedge K_2$ is consistent as well. But, since K_1 is a dictator, we have that $\Delta_\mu(\langle K_1, K_2\rangle) \equiv \Delta_\mu(\langle K_1\rangle)$. This leads to a contradiction. Weak Pareto and Strong Pareto are discussed in [8]. For Citizen's sovereignty take $E = \langle\{\varphi\}\rangle$, $\mu = \varphi$ and apply IC_2. For Participation take $w \in [\Delta_\mu(E) \wedge K]$. By IC_1, this implies that $w \in [\mu]$. We also have that $w \in [K]$, and from IC_2 it follows that $\Delta_\mu(\langle K\rangle) \equiv \bigwedge\langle K\rangle \wedge \mu$, hence $w \in [\Delta_\mu(\langle K\rangle)]$. This implies that $w \in [\Delta_\mu(E) \wedge \Delta_\mu(\langle K\rangle)]$, and by IC_5 we get that $w \in [\Delta_\mu(E \sqcup \langle K\rangle)]$. For Resolvability, take $K = \{\varphi\}$. By IC_0 it follows that $\varphi \models \mu$. Hence, $K \wedge \mu$ is consistent, and by IC_2 it follows that $\Delta_\mu(\langle K\rangle) \equiv \bigwedge\langle K\rangle \wedge \mu \equiv \varphi$. It follows that $\Delta_\mu(E) \wedge \Delta_\mu(\langle K\rangle)$ is consistent. The conclusion follows by using IC_5–IC_6. Consistency follows from repeated application of IC_5, and Homogeneity from repeated application of IC_5–IC_6. For Self-agreement first show, using IC_0, IC_1 and IC_2, that $\Delta_\mu(\langle\Delta_\mu(E)\rangle) \equiv \Delta_\mu(E)$. From this, together with the fact that $\Delta_\mu(E) \wedge \Delta_\mu(\langle\Delta_\mu(E)\rangle)$ is consistent, plus IC_5–IC_6, we get $\Delta_\mu(E \sqcup \langle\Delta_\mu(E)\rangle) \equiv \Delta_\mu(E) \wedge \Delta_\mu(\langle\Delta_\mu(E)\rangle) \equiv \Delta_\mu(E)$. □

The remaining properties require a different kind of analysis. Below we present a series of conditions on assignments which turn out to characterize several important remaining properties.

Definition 3. For an assignment on profiles, we define the following properties, for any $w, w_1, w_2 \in \mathcal{W}$, $K_1, \ldots, K_n \in \mathcal{K}$, $E, E_1, E_2, E_3 \in \mathcal{E}$, $\mathcal{M} \subseteq \mathcal{W}$ and transposition τ:[8]

($\mathsf{s}_{\mathsf{Neut}}$) If $w_1 \leq_E w_2$, then $\tau(w_1) \leq_{\tau(E)} \tau(w_2)$.
($\mathsf{s}_{\mathsf{Maj}}$) If $w_1 \leq_{\langle K_i \rangle} w_2$ for a majority of $i \in \{1, \ldots, n\}$, then $w_1 \leq_{\langle K_1, \ldots, K_n \rangle} w_2$.
($\mathsf{s}_{\mathsf{Con}}$) If $w \in \mathcal{M}$ is a weak Condorcet winner with respect to the preference profile $\langle \leq_{\langle K_1 \rangle}, \ldots, \leq_{\langle K_n \rangle} \rangle$ and $\mathcal{M}$ is the set of candidates, then $w \in \min_{\leq_{\langle K_1, \ldots, K_n \rangle}} \mathcal{M}$.
($\mathsf{s}_{\mathsf{Mon}}$) If $w_1 \leq_{E_1 \sqcup E_2} w_2$ and $w_1 \leq_{E_3} w_2$, then $w_1 \leq_{E_1 \sqcup E_3} w_2$.
($\mathsf{s}_{\mathsf{Rev}}$) If $w_1 <_E w_2$, then $w_2 <_{\neg E} w_1$.

We say that *property* s *of assignments characterizes property* P *of merging operators* if it holds that a merging operator satisfies P iff it is represented by an assignment satisfying s.

Theorem 3. *Properties* $\mathsf{s}_{\mathsf{Neut}}$, $\mathsf{s}_{\mathsf{Maj}}$, $\mathsf{s}_{\mathsf{Con}}$, $\mathsf{s}_{\mathsf{Mon}}$ *and* $\mathsf{s}_{\mathsf{Rev}}$ *characterize* Neutrality, SC-Majority, Condorcet's criterion, Monotonicity *and* Reversal symmetry, *respectively.*

Proof. For $\mathsf{s}_{\mathsf{Maj}}$, $\mathsf{s}_{\mathsf{Con}}$, $\mathsf{s}_{\mathsf{Mon}}$ and $\mathsf{s}_{\mathsf{Rev}}$ it is straightforward to check that they characterize their respective properties for operators. For $\mathsf{s}_{\mathsf{Neut}}$, suppose first that we have a neutral assignment and a merging operator Δ represented by it. We know that any renaming ρ is the product of n transpositions. We show that Δ satisfies Neutrality by induction on n. In the base case of $n = 0$, ρ is the identity renaming and the claim holds trivially. For the inductive step, we assume the claim holds for permutations of length n, and show that it holds for permutations of length $n+1$. Take, then, a renaming $\rho = \tau_1 \cdots \tau_n \tau_{n+1}$. By the inductive hypothesis, we know that $\tau_1 \cdots \tau_n(\Delta_\mu(E)) \equiv \Delta_{\tau_1 \cdots \tau_n(\mu)}(\tau_1 \cdots \tau_n(E))$. We apply τ_{n+1} to both sides. Using the results that for any $\varphi \in \mathcal{L}$ and transposition τ it holds that $[\tau(\varphi)] = \tau([\varphi])$ and $\tau(\tau(w)) = w$, and that for any $E \in \mathcal{E}$, $\mu \in \mathcal{L}$, it holds that $\tau(\Delta_\mu(E)) \equiv \Delta_{\tau(\mu)}(\tau(E))$, we derive the conclusion. Conversely, assume Δ is a merging operator that satisfies Neutrality but is represented by an assignment that does not satisfy $\mathsf{s}_{\mathsf{Neut}}$. Then there exists $E \in \mathcal{E}$, a transposition τ and $w_1, w_2 \in \mathcal{W}$ such that $w_1 \leq_E w_2$ and $\tau(w_2) <_{\tau(E)} \tau(w_1)$. Take $\mu \in \mathcal{L}$ such that $[\mu] = \{w_1, w_2\}$. We have that $w_1 \in [\Delta_\mu(E)]$ and hence $\tau(w_1) \in [\tau(\Delta_\mu(E))]$. On the other hand, $[\Delta_{\tau(\mu)}(\tau(E))] = \{w_2\}$. This shows that Δ is not neutral, which is a contradiction. □

Theorem 4. *None of the operators* Δ^{d_H} *and* Δ^{d_D} *satisfies* Monotonicity *or* Independence of clones. *The operators* Δ^{d_H} *do not satisfy* Entity resolution, Condorcet's criterion *and* Reversal symmetry, *but* Δ^{d_D} *does. Furthermore, there is no* IC *merging operator that satisfies* SC-Majority.

Proof. We provide here the relevant counter-examples. For Monotonicity take $K_1 = \{p \wedge q\}$, $K_2 = \{\neg q\}$, $K_3 = \{p\}$, $\mu = p$ and $E_1 = \langle K_1 \rangle$, $E_2 = \langle K_2 \rangle$, $E_3 = \langle K_3 \rangle$. We get that $\Delta_\mu^{d_H}(E_1 \sqcup E_2) \equiv \Delta_\mu^{d_D}(E_1 \sqcup E_2) \equiv \{p\}$, $\Delta_\mu^{d_H}(E_3) \equiv \Delta_\mu^{d_D}(E_3) \equiv \{p\}$ and $\Delta_\mu^{d_H}(E_1 \sqcup E_3) \equiv \Delta_\mu^{d_D}(E_1 \sqcup E_3) \equiv \{p \wedge q\}$. For Independence of clones take $E = \langle K_1, K_2 \rangle$, $K_1 = \{p\}$ and $K_2 = \{q\}$, $\mu = p \vee q$ and add a clone r of p. Then $\Delta_\mu^{d_H}(E) \equiv \{p \wedge q\}$. In the new setup, we get that $\Delta_\mu^{d_H}(E^{p,\mathcal{Q}}) \equiv \{p \wedge q \wedge r\}$. For Entity resolution and Δ^{d_H}, take $E = \langle K_1, K_2, K_3 \rangle$ with $K_1 = \{p \wedge \neg q \wedge r\}$, $K_2 = \{p \wedge \neg r\}$, $K_3 = \{\neg p \wedge \neg q\}$ and $\mu = \top$. We get $\Delta^{d_H}_{\mu \wedge (p \leftrightarrow q)}(\langle K_1, K_2 \rangle) \equiv \{\neg p \wedge \neg q\}$, while $E^{p/q} = \langle \{q \wedge \neg r\}, \{\neg q\} \rangle$ and $\Delta^{d_H}_{\mu^{p/q}}(E^{p/q}) \equiv \{\neg r\}$. For Entity resolution and Δ^{d_D}, take $w \in [\Delta^{d_D}_{\mu \wedge (p \leftrightarrow q)}(E)]$. We have that if $w \in [K_i]$ and $K_i^{p/q}$ is consistent, then $w \in [K_i^{p/q}]$. This implies that the number of 0's in w's vector of scores in $\leq_{E^{p/q}}$ is equal to the number of 0's in w's vector of scores in $\leq_{E^{p/q}}$. If $K_i^{p/q}$ is inconsistent, this is because K_i implies either $p \wedge \neg q$ or $\neg p \wedge q$, and thus w cannot be a model of K_i; hence removing K_i can only decrease w's final score. For Condorcet's criterion and $\Delta^{d_H,\Sigma}$ or $\Delta^{d_H,GMAX}$, take $E = \langle K_1, K_2, K_3 \rangle$, $K_1 = \{p \wedge q \wedge r\}$, $K_2 = K_3 = \{\neg p\}$, $\mu = (\neg p \wedge \neg q \wedge \neg r) \vee (p \wedge q \wedge r)$. Then $[K_1] = \{111\}$, $[K_2] = [K_3] = \{000, 001, 010, 011\}$ and $[\mu] = \{000, 111\}$. We have that $\varphi = \neg p \wedge \neg q \wedge \neg r$ is the only weak Condorcet winner with respect to E and μ, but $\Delta_\mu^{d_H,\Sigma}(E) \equiv \Delta_\mu^{d_H,GMAX}(E) \equiv \{p \wedge q \wedge r\}$. For $\Delta^{d_H,GMIN}$, take $E = \langle K_1, K_2, K_3 \rangle$, $K_1 = \{\neg p \wedge \neg q \wedge r\}$, $K_2 = \{\neg p \wedge q \wedge \neg r\}$, $K_3 = \{p \wedge q \wedge r\}$ and μ from before. Then the weak Condorcet winner with respect to E and μ is $\neg p \wedge \neg q \wedge \neg r$, but $\Delta_\mu^{d_H,GMIN}(E) \equiv \{p \wedge q \wedge r\}$. It is straightforward to check that d_D together with any aggregation function generates an assignment that satisfies $\mathsf{s}_{\mathsf{Con}}$. Together with Theorem 3 and our observation that a weak Condorcet winner φ with respect to E and μ corresponds to $[\varphi]$ being a weak Condorcet winner in the voting profile $\langle \leq_{\langle K_1 \rangle} \ldots, \leq_{\langle K_n \rangle} \rangle$ restricted to $[\mu]$, we get that Δ^{d_D} satisfies Condorcet's criterion. For Reversal symmetry and d_H, take $K_1 = \{p \rightarrow q\}$, $K_2 = \{p \wedge \neg q\}$, $\mu = \neg p$ and $E = \langle K_1, K_2 \rangle$. We get that $\Delta_\mu^{d_H}(E) \equiv \Delta_\mu^{d_H}(\neg E) \equiv \{\neg p \wedge \neg q\}$. It is straightforward to check that d_D with any aggregation function generates an assignment that satisfies $\mathsf{s}_{\mathsf{Rev}}$, thus Δ^{d_D} satisfies Reversal symmetry. To see why the IC postulates and SC-Majority are incompatible, suppose there is an IC merging operator which satisfies SC-Majority. Take $[\mu] = \{w_1, w_2\}$ and $[K_1] = \{w_1, w_2\}$, $[K_2] = \{w_1, w_2\}$, $[K_3] = \{w_1\}$. By SC-Majority we get that $\{w_1, w_2\} \subseteq [\Delta_\mu(\langle K_1, K_2, K_3 \rangle)]$. However, by IC_2 we get that $[\Delta_\mu(\langle K_1, K_2, K_3 \rangle)] = \{w_1\}$. □

Finally, we show that Neutrality is not implied by the IC postulates, even though Δ^{d_H} and Δ^{d_D} satisfy it. To do so, we first provide a characterization of Neutrality in terms of a corresponding property for distance based operators. We call a pseudo-distance d *neutral* if for any transposition τ and $w_1, w_2 \in \mathcal{W}$, it holds that $d(w_1, w_2) = d(\tau(w_1), \tau(w_2))$. The characterization is then captured by the following result.

Theorem 5. *For any pseudo-distance* d *and aggregation function* f*, a merging operator* $\Delta^{d,f}$ *satisfies* Neutrality *if and only if* d *is neutral.*

Proof. For one direction of the proof, take an assignment generated using a neutral distance d and an aggregation function f. First we show that for any $w \in \mathcal{W}$, $K \in \mathcal{K}$ and transposition τ, it holds that $d(w, K) = d(\tau(w), \tau(K))$. Take $w' \in [K]$ such that $d(w, K) = d(w, w')$. Since τ is neutral, we get that $d(w, w') = d(\tau(w), \tau(w'))$. We have that $[\tau(K)] = \tau([K])$, and thus $\tau(w') \in [\tau(K)]$. We show now that $\tau(w')$ is at a minimal distance from $\tau(w)$ among the models of $\tau(K)$. Take, then, $\tau(w'') \in [\tau(K)]$, with $w'' \in [K]$. We have that $d(w, w') \leq d(w, w'')$, and since d is neutral it follows that $d(\tau(w), \tau(w')) \leq d(\tau(w), \tau(w''))$. Hence $d(\tau(w), \tau(K)) = d(\tau(w), \tau(w')) = d(w, w')$. From this we immediately derive that for any $E \in \mathcal{E}$, $w \in \mathcal{W}$ and neutral transposition τ, it holds that $d(w, E) = d(\tau(w), \tau(E))$. Thus, if $d(w_1, E) \leq d(w_2, E)$, then $d(\tau(w_1), \tau(E)) \leq d(\tau(w_2, \tau(E)))$, for any $w_1, w_2 \in \mathcal{W}$. It follows that if $w_1 \leq_E w_2$ then $\tau(w_1) \leq_{\tau(E)} \tau(w_2)$, and therefore the assignment satisfies $\mathsf{s}_{\mathsf{Neut}}$. By Theorem 3 this implies $\Delta^{d,f}$ is neutral.

[8] We remind the reader that transpositions applied to formulas swap exactly two atoms among each other and applied to interpretations they swap the corresponding bits in the bit-vector representation.

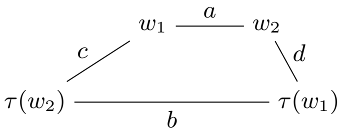

Figure 1: Distances between w_1, w_2, $\tau(w_1)$ and $\tau(w_2)$

Conversely, assume $\Delta^{d,f}$ satisfies Neutrality, but d is not neutral. Then there must be $w_1, w_2 \in \mathcal{W}$ and a transposition τ such that $d(w_1, w_2) \neq d(\tau(w_1), \tau(w_2))$. We show now that, by taking $\rho = \tau$, we can always find some profile E and constraint μ such that $\Delta^{d,f}$ does not satisfy Neutrality. Notice first that it is not possible to have $\tau(w_1) = w_1$ and $\tau(w_2) = w_2$, as this contradicts our assumption that $d(w_1, w_2) \neq d(\tau(w_1), \tau(w_2))$. In the following we analyze the remaining cases, keeping in mind that $\tau(\tau(w)) = w$ and that $[\tau(K)] = \tau([K])$. Let us denote $d(w_1, w_2) = a$, $d(\tau(w_1), \tau(w_2)) = b$, $d(w_1, \tau(w_2)) = c$ and $d(\tau(w_1), w_2) = d$ (see Figure 1). Without loss of generality, we assume that $a < b$. *Case 1.* $\tau(w_1) = w_1$, $\tau(w_2) \neq w_2$. Take K and μ such that $[K] = \{w_1\}$ and $[\mu] = \{w_2, \tau(w_2)\}$. We have that $[\tau(K)] = \tau([K]) = \{\tau(w_1)\} = \{w_1\}$, and $[\tau(\mu)] = [\mu]$. Then by $b = c$ we obtain $[\Delta^{d,f}_{\mu}(\langle K\rangle)] = [\Delta^{d,f}_{\tau(\mu)}(\tau(\langle K\rangle))] = \{w_2\}$. This shows that $\Delta^{d,f}$ is not neutral, since $[\tau(\Delta^{d,f}_{\mu}(\langle K\rangle))] = \tau([\Delta^{d,f}_{\mu}(\langle K\rangle)]) = \{\tau(w_2)\}$. *Case 2.* $\tau(w_1) \neq w_1$, $\tau(w_2) = w_2$. Analogous to Case 1. *Case 3.* $\tau(w_1) \neq w_1$, $\tau(w_2) \neq w_2$. For this case we have to analyze the relationship between a, b, c and d. *Case 3.1.* $\min\{a, c\} < \min\{b, d\}$ or $\min\{b, d\} < \min\{a, c\}$. Take $[K] = \{w_2, \tau(w_2)\}$ and $[\mu] = \{w_1, \tau(w_1)\}$. Clearly, $[\tau(K)] = [K]$ and $[\tau(\mu)] = [\mu]$. In this case we have that $d(w_1, K) = \min\{a, c\}$ and $d(\tau(w_1), K) = \min\{b, d\}$. Then $[\Delta^{d,f}_{\mu}(\langle K\rangle)]$ will consist of exactly one interpretation out of $\{w_1, \tau(w_1)\}$, call it w (see Table 4). But this shows that $\Delta^{d,f}$ cannot be neutral, because we will get the

w	$\{w_2, \tau(w_2)\}$
w_1	$\min\{a, c\}$
$\tau(w_1)$	$\min\{b, d\}$

Table 4: $\min\{a, c\} \neq \min\{b, d\}$

same result of $\{w\}$ for $[\Delta^{d,f}_{\tau(\mu)}(\tau(\langle K\rangle))]$, while $[\tau(\Delta^{d,f}_{\mu}(\langle K\rangle))] = \{\tau(w)\}$. *Case 3.2.* $\min\{a, c\} = \min\{b, d\}$. Here we analyze two sub-cases, but the reasoning follows the same lines as in the previous cases. *Case 3.2.1.* $a \leq c$, $d \leq b$, $a = d$. Take K and μ such that $[K] = \{w_1\}$ and $[\mu] = \{w_2, \tau(w_2)\}$. Then $[\tau(K)] = \{\tau(w_1)\}$ and $[\tau(\mu)] = [\mu]$ and we get that $[\Delta^{d,f}_{\mu}(\langle K\rangle)] = [\Delta^{d,f}_{\tau(\mu)}(\tau(\langle K\rangle))] = \{w_2\}$, whereas $[\tau(\Delta^{d,f}_{\mu}(\langle K\rangle))] = \{\tau(w_2)\}$. *Case 3.2.2.* $c \leq a$, $d \leq b$, $c = d$. Take K_1, K_2 and μ such that $[K_1] = \{w_2\}$, $[K_2] = \{\tau(w_2)\}$, $[\mu] = \{w_1, \tau(w_1)\}$. Then $[\tau(K_1)] = \{\tau(w_2)\}$, $[\tau(K_2)] = \{w_2\}$ and $[\tau(\mu)] = [\mu]$. Notice, now, that from $c = d$, $a < b$ and properties (i) and (iv) of f as an aggregation function (see Section 2), it follows that $f(a, c) < f(d, b)$. Consequently, $[\Delta^{d,f}_{\mu}(\langle K_1, K_2\rangle)] = [\Delta^{d,f}_{\tau(\mu)}(\tau(\langle K_1, K_2\rangle))] = \{w_1\}$. However, $[\tau(\Delta^{d,f}_{\mu}(\langle K_1, K_2\rangle))] = \{\tau(w_1)\}$. Thus, $\Delta^{d,f}$ is not neutral. □

Using Theorem 5, we can now state our last result.

Theorem 6. Neutrality *does not follow from the* IC *postulates, but* Δ^{d_H} *and* Δ^{d_D} *satisfy it.*

Proof. It is straightforward to check that d_H and d_D are neutral, hence by Theorem 5 it follows that Δ^{d_H} and Δ^{d_D} satisfy Neutrality. However Neutrality is not guaranteed by the IC postulates: there exist distance-based operators satisfying the IC postulates where the distance d is nonetheless not neutral. One such example is a merging operator for the Horn fragment of propositional logic based on a custom-defined distance d_S [11]. In a three letter alphabet the definition of d_S specifies that $d_S(000, 001) = 1$ and $d_S(000, 010) = 2$. Thus, d_S is not neutral, which can be seen by considering the transposition that swaps the second and third bits among themselves. Nonetheless, $\Delta^{d_S,\Sigma}$ satisfies the IC postulates [11]. □

We conclude by a few comments on our results. First, notice that Neutrality for distance-based operators depends only on the distance used and not on the aggregation function. Concerning the connection between our results and social choice, at first glance it might look disappointing that there is no IC operator satisfying SC-Majority, but it should be kept in mind that there are also important voting rules (e.g., Borda) which do not satisfy this property. Furthermore, it is a positive result that Participation holds for all IC operators, as this is not the case for important voting rules (e.g., Copeland, Dodgson and Young). Having Participation removes an agent's incentive for strategizing about whether to cast a vote. Also, it is not overly surprising that Independence of clones does not hold for all IC operators as it does not hold for many voting rules either (e.g., Plurality, Borda, Copeland and Dodgson). Finally, the result on Consistency is notable as this property does not hold for several important voting rules (e.g., Copeland, Dodgson and Young).

Note that we consider the strongest setting, where the constraint μ is unrestricted and properties have to hold for any μ and any profile. The cases when either domain restrictions are imposed on μ, or $\mu \equiv \top$, remain to be explored. Some of the proofs will carry over, whereas several results will have to be revisited.

5 Conclusion

In this work we have investigated eighteen desirable properties for belief merging operators, fourteen of which are newly formulated using insights from voting theory. We show that some follow from the IC postulates, some can never be satisfied by an IC operator, whereas others are only satisfied by certain IC operators. For properties of the last case, we additionally verified which of the standard operators satisfy them. If a property already follows from the postulates this is good news; if it does not, this shows that special care is needed when designing tailor-made operators. The properties proposed in this work are, however, to be seen as a first step on a long path and can certainly be refined and extended.

Likewise, there are quite a number of possible directions for future work. A natural step is to perform an extensive classification which is not limited to standard operators. One could also have a closer look at operators for fragments such as Horn, whenever a property is not satisfied in the general setting. Furthermore, for each considered property it is enticing to come up with a representation theorem for the setting where the IC postulates are extended by this property. Certainly also the relations between the properties studied in this work deserve a closer investigation. In particular, it would be interesting to come up with (IC) merging operators satisfying a maximal number of properties and to complement these results with impossibility theorems for the remaining cases. Also, as discussed above, the role of the constraint formula μ deserves a closer investigation, in which domain restrictions of μ are considered. Last but not least, we plan to explore the relation between judgment aggregation and belief merging—for the general case this relation was recently studied by Everaere, Konieczny, and Marquis [9]—with a special focus on devising new suitable properties.

Acknowledgments

This work was supported by the Austrian Science Fund (FWF) under grants P25518 and P25521, and the German Research Foundation (DFG): ER 738/2-1.

REFERENCES

[1] *Handbook of Computational Social Choice*, eds., Felix Brandt, Vincent Conitzer, Ulle Endriss, Jèrôme Lang, and Ariel D. Procaccia, Cambridge University Press, 2016.
[2] Yann Chevaleyre, Ulle Endriss, Jérôme Lang, and Nicolas Maudet, 'Preference handling in combinatorial domains: From AI to social choice', *AI Magazine*, **29**(4), 37–46, (2008).
[3] Samir Chopra, Aditya K. Ghose, and Thomas A. Meyer, 'Social choice theory, belief merging, and strategy-proofness', *Information Fusion*, **7**(1), 61–79, (2006).
[4] Amílcar Mata Díaz and Ramón Pino Pérez, 'Logic-based fusion of complex epistemic states', in *Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 11th European Conference, ECSQARU. Proceedings*, volume 6717 of *Lecture Notes in Computer Science*, pp. 398–409. Springer, (2011).
[5] Daniel Eckert and Gabriella Pigozzi, 'Belief merging, judgment aggregation and some links with social choice theory', in *Belief Change in Rational Agents: Perspectives from Artificial Intelligence, Philosophy, and Economics*, volume 05321 of *Dagstuhl Seminar Proceedings*. IBFI, Schloss Dagstuhl, (2005).
[6] Ulle Endriss, 'Applications of logic in social choice theory - (invited talk)', in *Computational Logic in Multi-Agent Systems - 12th International Workshop, CLIMA XII. Proceedings*, volume 6814 of *Lecture Notes in Computer Science*, pp. 88–91. Springer, (2011).
[7] Patricia Everaere, Sébastien Konieczny, and Pierre Marquis, 'The strategy-proofness landscape of merging', *Journal of Artificial Intelligence Research (JAIR)*, **28**, 49–105, (2007).
[8] Patricia Everaere, Sébastien Konieczny, and Pierre Marquis, 'On egalitarian belief merging', in *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourteenth International Conference, KR 2014*. AAAI Press, (2014).
[9] Patricia Everaere, Sébastien Konieczny, and Pierre Marquis, 'Belief merging versus judgment aggregation', in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015*, pp. 999–1007. ACM, (2015).
[10] Dov M. Gabbay, Odinaldo Rodrigues, and Gabriella Pigozzi, 'Connections between belief revision, belief merging and social choice', *Journal of Logic and Computation*, **19**(3), 445–446, (2009).
[11] Adrian Haret, Stefan Rümmele, and Stefan Woltran, 'Merging in the Horn fragment', in *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015*, pp. 3041–3047. AAAI Press, (2015).
[12] Sébastien Konieczny and Ramón Pino Pérez, 'Propositional belief base merging or how to merge beliefs/goals coming from several sources and some links with social choice theory', *European Journal of Operational Research*, **160**(3), 785–802, (2005).
[13] Sébastien Konieczny and Ramón Pino Pérez, 'Merging information under constraints: A logical framework', *Journal of Logic and Computation*, **12**(5), 773–808, (2002).
[14] Sébastien Konieczny and Ramón Pino Pérez, 'Logic Based Merging', *Journal of Philosophical Logic*, **40**(2), 239–270, (2011).
[15] Pedrito Maynard-Zhang and Daniel J. Lehmann, 'Representing and aggregating conflicting beliefs', *Journal of Artificial Intelligence Research (JAIR)*, **19**, 155–203, (2003).
[16] *Economics and Computation*, ed., Jörg Rothe, Springer, 2015.
[17] Nicolaus Tideman, *Collective Decisions and Voting: The Potential for Public Choice*, Ashgate, 2006.