The Epistemic View of Belief Merging: Can We Track the Truth?

Patricia Everaere
LIFL-CNRS
USTL, France
patricia.everaere@univ-lille1.fr

Sébastien Konieczny
CRIL-CNRS
Université d’Artois, France
konieczny@cril.fr

Pierre Marquis
CRIL-CNRS
Université d’Artois, France
marquis@cril.univ-artois.fr

Abstract
Belief merging is often described as the process of defining a belief base which best represents the beliefs of a group of agents (a profile of belief bases). The resulting belief base can be viewed as a synthesis of the input profile. In this paper another view of what belief merging aims at is considered: the epistemic view. Under this view the purpose of belief merging is to best approximate what is the true state of the world. We study generalizations of Condorcet’s Jury Theorem from the belief merging perspective. Roughly we show that if we merge the beliefs of sufficiently many reliable agents then we can ensure to identify the true state of the world. We show that some merging operators from the literature are suited to the truth tracking issue.

1 Introduction
In many areas of computer science, including distributed databases and multi-agent systems, one needs to synthesize pieces of information from several information sources. What makes this problem difficult is, among other things, that the information sources typically contradict each other. When the available pieces of information are beliefs represented in propositional logic, this problem is called (propositional) belief merging. Many different belief merging operators have been pointed out so far. In this paper we focus on the purely logical case, i.e., we assume that belief bases are sets of propositional formulae (see e.g. [Baral et al., 1991; Revesz, 1997; Konieczny and Pino Pérez, 2002; Konieczny et al., 2004; Everaere et al., 2005]). Other operators have been provided in more general settings, such as weighted logics (possibilistic logic or settings based on ordinal conditional functions) [Benferhat et al., 2002; Meyer, 2001; Benferhat et al., 2007], which prove useful when more information are available (especially, when all the pieces of belief are not equally certain). In these more general settings, the merging problem becomes close to preorder (preference) aggregation, as studied in social choice theory [Arrow, 1963; Pini et al., 2005].

Logical properties of merging operators have been investigated in several works [Revesz, 1997; Liberatore and Schaerf, 1998; Konieczny and Pino Pérez, 2002]. In [Konieczny and Pino Pérez, 2002] a set of logical properties have been put forward to characterize the family of IC (Integrity Constraints) merging operators. IC merging operators have been advocated to be suited to both belief merging and goal merging. Even if it might look strange at a first glance that very different concepts, such as goals and beliefs, can be handled in the same way with respect to aggregation, the adequacy of IC merging operators to propositional merging (whatever goals or beliefs are to be merged) has not been challenged so far. This makes sense since in both cases merging aims at synthesizing the information represented in the given profile of propositional bases.

In this paper we introduce a new point of view about belief merging, that goes beyond the usual synthesis view: the epistemic view.

Synthesis View: Under the synthesis view, as explained above, belief merging aims at characterizing a base which best represents the beliefs of the input profile. This is the view adopted in previous belief merging works.

Epistemic View: Under the epistemic view, the purpose of a belief merging process is to best approximate what is the true state of the world.

In the general case, no agent has a perfect view of the real world, her beliefs are pervaded with uncertainty:

- An agent typically does not know which one of the models of her base represents the true state of the world,
- She is not even ensured that the true state of the world is really among the models of her base.\(^1\)

Belief merging under the epistemic view can be considered as a way to circumvent such an uncertainty at the group level. Interestingly, the truth tracking issue is a way to discriminate belief merging from goal merging. Indeed, while truth tracking can be expected in many cases when beliefs are to be merged, the concept of truth tracking is meaningless when goals are considered. Obviously enough, there is no notion of

\(^1\)If one supposes that the agent is ensured that the true state of the world is a model of her belief base, then one talks about “knowledge” – this assumption is the only difference between belief and knowledge – and knowledge merging is not so interesting, since the only sensible knowledge merging operator is obviously conjunction.
“right goal” which would be analogous to the true state of the world in the goal merging setting.

Now, the intuition underlying the epistemic view of belief merging is that if agents are independent and reliable then we could expect from a belief merging operator to be able to identify the true state of the world when listening sufficiently many agents.

The problem of truth tracking has been studied for centuries in social choice and in political science, in order to justify the foundations of democratic elections or of decisions made by jury trials. The main theoretical result here is Condorcet’s Jury Theorem [Condorcet, 1785]. This theorem states that if a jury is composed of reliable and independent individuals, and if they have to find the true answer to a yes/no question, then the probability that the decision made by the jury is the good one tends to 1 as the size of the jury tends to infinity.

In this paper we formalize the truth tracking issue from a belief merging perspective. We show that some belief merging operators can be used to identify the true state of the world by considering sufficiently many reliable, homogeneous and independent agents. More precisely, the contribution of the paper is mainly as follows: we present a generalization of Condorcet’s Jury Theorem under uncertainty (i.e., when each base may have several models). We introduce a Truth Tracking (TT) postulate, present some belief merging operators satisfying it and show that TT is independent of the (conjunction of the) IC postulates. As a consequence, we conclude that some, but not all, merging operators from the literature are good for the truth tracking issue. We also provide experimental results in order to investigate the convergence speed of truth tracking for the belief merging operator \( \Delta_{dD,\Sigma} \). In most cases the number of agents to be considered for ensuring that the merged base identifies the true state of the world with high probability is not so huge. This shows the practical feasibility of truth tracking using \( \Delta_{dD,\Sigma} \).

The layout of the paper is as follows: we give some formal preliminaries in Section 2. We recall Condorcet’s Jury Theorem, and some of its generalizations in Section 3. In Section 4 we point out new generalizations of Condorcet’s Jury Theorem suited to the belief merging perspective. In Section 5 we show that IC postulates and TT are logically independent. In Section 6, we present the empirical results we obtained and we discuss them. Finally we conclude in Section 7.

## 2 Preliminaries

We consider a propositional language \( \mathcal{L} \) defined from a finite set of propositional variables \( \mathcal{P} \) and the usual connectives.

For any subset \( e \) of \( \mathcal{P} \), \( |e| \) denotes the number of elements of \( e \). An interpretation (or state of the world) is a total function from \( \mathcal{P} \) to \{0, 1\}. The set of all interpretations is noted \( \mathcal{W} \). The true state of the world is noted \( \omega^* \). An interpretation \( \omega \) is a model of a formula \( \phi \in \mathcal{L} \) if and only if it makes it true in the usual truth functional way. \( \langle \cdot \rangle \) denotes the set of models of formula \( \phi \), i.e., \( \langle \phi \rangle = \{ \omega \in \mathcal{W} \mid \omega \models \phi \} \).

A base \( K \) denotes the set of beliefs of an agent, it is a finite set of propositional formulae, interpreted conjunctively. We identify \( K \) with the conjunction of its elements. Basically, a base \( K \) represents a set \([K]\) of states of the world.

A profile \( E \) denotes the beliefs from the group of \( n \) agents that are involved in the merging process. In this paper agents express sometimes only a single possible world. In this case a profile is a vector of complete bases. In order to avoid heavy notations, we assimilate each complete base with its model and write such profiles as \( E_n = \{\omega_1, \ldots, \omega_n\} \). Elsewhere agents express sets of possible worlds, hence \( E \) is represented as a vector of bases \( E = \{K_1, \ldots, K_n\} \), as usual in propositional merging.

In the following, agents \( 1, \ldots, n \) are identified with the corresponding belief bases \( K_1, \ldots, K_n \). When unknown, each \( K_i \) can also be viewed as a discrete random variable, ranging over \( 2^\mathcal{W} \) (or \( \mathcal{W} \) when each agent has to report a complete base). The true state of the world is usually unknown as well so it is also viewed as a random variable \( \mathcal{W}^* \), ranging over \( \mathcal{W} \).

Two important notions about sets of agents will be considered in the following: independence and homogeneity.

Agents are independent if knowing the true state of the world and a state of the world reported by any other agent does not give any further information on the states of the world given by an agent \( i \) (this means that the agents choices are independent conditionally to the true state of the world in a standard Bayesian way [Pearl, 2000]). Formally, agents \( 1, \ldots, n \) are said to be independent if \( \forall \omega, \omega_1, \ldots, \omega_n \in \mathcal{W} \):

\[
P(\bigwedge_{i=1}^n \omega_i = K_i \mid \mathcal{W}^* = \omega) = \prod_{i=1}^n P(\omega_i = K_i \mid \mathcal{W}^* = \omega).
\]

Obviously, when agents report complete bases, the formal definition of independence can be stated as follows: agents \( 1, \ldots, n \) are independent if \( \forall \omega, \omega_1, \ldots, \omega_n \in \mathcal{W} : P(\bigwedge_{i=1}^n K_i = \{\omega_i\}) \mid \mathcal{W}^* = \omega = \prod_{i=1}^n P(K_i = \{\omega_i\}) \mid \mathcal{W}^* = \omega \).

Agents \( 1, \ldots, n \) are homogenous if for every world \( \omega_i \in \mathcal{W} \), the probability \( P(\omega_i = K_i) \) that \( \omega_i \) is a model of the base \( K_i \) of the profile \( E \) is the same for all the agents \( i \in 1, \ldots, n \) of the set. In particular, the real world \( \omega^* \) has the same probability to appear as a model for each agent.

## 3 Condorcet’s Jury Theorem and Extensions

We consider a profile \( E_n \) of \( n \) agents where each agent \( i \) votes for an alternative, let us say a state of the world \( \omega_i \in \mathcal{W} \). Among the possible states of the world is the true one \( \omega^* \).

The hypotheses used in Condorcet’s Jury Theorem are that agents are both independent and reliable. Since several notions of reliability will be considered in the following, we call the first one R1-reliability:

- The R1-reliability \( p_i \) of an agent \( i \) is the probability that \( i \) gives the true state of the world, i.e.,
  \[ p_i = P(K_i = \{\omega^*\}) \].

- An agent \( i \) is R1-reliable if her R1-reliability is strictly greater than 0.5. (R1)

The majority rule simply returns as result the interpretation which receives a strict majority of votes. Formally, let us first
define the notion of score of a world w.r.t a profile of complete bases:
\[ s(\omega) = |\{\omega_i \in E_{c} \text{ s.t. } \omega_i = \omega\}|. \]

**Definition 1 (Majority)** Given a profile \( E_c \) of \( n \) complete bases, the majority rule \( m \) is defined as:
\[ m(E_c) = \omega \text{ if } s(\omega) > n/2. \]

We are now ready to recall Condorcet’s Jury Theorem. In this theorem, only two alternatives are considered so that each agent votes for one of them:

**Theorem 1 ([Condorcet, 1785])** Consider two possible states of the world \( W = \{\omega, \omega^*\} \) and a profile \( E_c \) of complete bases from a set of \( n \) independent and R1-reliable agents sharing the same R1-reliability. The probability that the majority rule on this profile returns the true state of the world \( \omega^* \) tends to 1 as \( n \) tends to infinity, i.e.:
\[ P(m(E_c) = \omega^*) \xrightarrow{n \to \infty} 1. \]

This theorem is a consequence of the (weak) law of large numbers. Roughly, it states that if the individuals in a jury are sufficiently reliable (they perform better than pure randomizers) and independent, then the probability that the jury makes the right decision tends to 1 when the size of the jury tends to infinity.

Clearly enough, the assumptions used in this theorem are quite strong. First, usually agents in a jury are not fully independent: they often have a similar background, listen the same opinion leaders, etc. Furthermore, in general, all the agents do not have exactly the same reliability: there are usually agents more competent than others. Interestingly, some extensions of Condorcet’s Jury Theorem show that strong assumptions can be relaxed without questioning the conclusion. Thus, the theorem still holds when the opinions of the individuals are not independent [Estlund, 1994]. And as far as reliability is concerned, it is enough to assume that the mean reliability of the individuals is above 0.5 [Owen et al., 1989].

A further limitation of Condorcet’s Jury Theorem is that it considers only two alternatives. A recent result by [List and Goodin, 2001] allows to extend the theorem to any finite number of options. In order to present this result, we first need to recall the definition of the plurality rule:

**Definition 2 (Plurality)** Given a profile \( E_c \) of complete bases, the plurality rule \( pl \) is defined as:
\[ pl(E_c) = \{\omega \text{ s.t. } \forall \omega' \in W \text{ s.t. } s(\omega') \geq s(\omega')\} \]

The reliability assumption \((R1)\) has to be extended to more than two alternatives. List and Goodin [List and Goodin, 2001] define the following notion of reliability; consider \( k \) possible states of the world \( W = \{\omega_1, \ldots, \omega_k, \omega^*\} \):

- An agent is \( R2 \)-reliable if the probability that she votes for \( \omega^* \) is strictly greater than the probability that she votes for another world. \( (R2) \)

List and Goodin showed that:

**Theorem 2 ([List and Goodin, 2001])** Consider \( k \) possible states of the world \( W = \{\omega_1, \ldots, \omega_{k-1}, \omega^*\} \) and a profile \( E_c \) of complete bases from a set of \( n \) independent, homogeneous, and \( R2 \)-reliable agents. The probability that the plurality rule on this profile returns the true state of the world \( \omega^* \) tends to 1 as \( n \) tends to infinity, i.e.:
\[ P(pl(E_c) = \{\omega^*\}) \xrightarrow{n \to \infty} 1. \]

This theorem is a generalization of Condorcet’s Jury Theorem. Observe that the plurality rule is used (not the majority rule) and that the reliability assumption only requires that the probability of voting for the true state of the world is strictly greater than the probability of voting for another world, so that the probability of voting for the true state of the world can be less than 0.5.

It is also interesting to observe that the homogeneity assumption is not used explicitly in Condorcet’s Jury Theorem. However, it is implicitly there, just because \( R1 \)-reliability implies homogeneity when only two alternatives are considered (the probability that any agent chooses a world different from \( \omega^* \) is \( 1-p \) if \( p \) is the agents’ \( R1 \)-reliability). Thus, when considering only two states of the world, the hypotheses used in List-Goodin’s theorem are equivalent to the ones used in Condorcet’s Jury Theorem, so that the two theorems are identical in this case, as expected.

See [List and Goodin, 2001] for more discussion on their theorem and its philosophical consequences, and for a discussion about Condorcet’s Jury Theorem.

## 4 A Jury Theorem under Uncertainty

In all these previous works around Condorcet’s Jury Theorem, agents are supposed to vote for a unique alternative. This makes them inadequate for our purpose since in belief merging, agents typically give belief bases having several models (and imposing agents to give complete belief bases would be very restrictive since it would deny that the agents’ beliefs can be uncertain). Thus, from now on, we assume that each agent \( i \) gives a belief base \( K_i \) which may have several models taken from a finite set \( W = \{\omega_1, \ldots, \omega_{k-1}, \omega^*\} \).

So let us show how the Jury Theorem can be extended to consider the case when each agent may vote for several alternatives. We first need to define a notion of agent reliability suited to this situation:

- The \( R3 \)-reliability \( p_i \) of an agent \( i \) is the probability that the true state of the world \( \omega^* \) is among the models of her belief base \( K_i \), i.e., \( p_i = P(\omega^* \models K_i) \).
- An agent is \( R3 \)-reliable if \( p_i > 0.5 \). \( (R3) \)

Similarly, the notion of score of a world has to be extended in the following way:
\[ s_\omega(\omega) = |\{K_i \in E \text{ s.t. } \omega \models K_i\}|. \]

Then it is possible to state the following result:

\(^2\)List and Goodin proposed a notion of reliability which encompasses both our \( R2 \)-reliability and homogeneity.
Proposition 1 Consider a real $p^* \in [0,1]$ and a profile $E$ from a set of $n$ independent agents which have the same R3-reliability $p > p^*$. The probability that the score of the true state of the world exceeds $p^* n$ tends to 1 when $n$ tends to infinity, i.e.,

$$P(s_\omega(\omega^*) > p^* n) \xrightarrow{n \to \infty} 1.$$  

Proof: For a given agent $i$, the probability that $\omega^* \models K_i$ is $p$; so the probability that $\omega^* \not\models K_i$ is $q = 1 - p$. Assuming that the agents are independent, the probability that $\omega^*$ is a model of $k$ out of the $n$ bases is $\binom{n}{k} p^k q^{n-k}$. The proof is then derived using the Bienaime-Tchebitcheff theorem: let $v$ denote a random variable with mean $m$ and variance $\sigma^2$. Given $t$, a nonnegative real number, the probability that $v$ deviates from $m$ by more than $t \sigma$ is smaller than $\frac{1}{t^2}$:

$$P(|v - m| \geq t \sigma) < \frac{1}{t^2}. \quad (1)$$

Let $v_i$ denote a random variable equals to 1 if $\omega^* \models K_i$, and to 0 otherwise. As each $v_i$ follows a Bernoulli distribution with a success probability $p$, the mean and variance of $v_i$ correspond to $p$ and $pq$, respectively. So the number of bases having $\omega^*$ as model is equal to $v = \sum_{i=1}^n v_i$. As $v$ follows a binomial distribution with parameters $n$ and $p$, the mean of $v$ corresponds to $m = np$, and its variance is equal to $\sigma^2 = npq$, due to the independence assumption.

The probability that the number of bases having $\omega^*$ as model is less than $np^*$ is given by $P(v \leq p^* n)$. We have $|v - m| \geq m - p^* n$ if and only if $v - m \geq m - p^* n \geq 0$ or (exclusive) $v - m \leq p^* n - m < 0$. So $P(|v - m| \geq m - p^* n) = P(v - m \geq m - p^* n \geq 0) + P(v - m \leq p^* n - m < 0)$.

Then $P(|v - m| \geq m - p^* n) \geq P(v - m \leq p^* n - m < 0)$. Since $v - m \leq p^* n - m < 0$ if and only if $v \leq p^* n < m$, we get:

$$P(v \leq p^* n < m) \leq P(|v - m| \geq m - p^* n).$$

With $t = \frac{m - p^* n}{\sigma}$ ($t > 0$ if $p > p^*$), and using inequality (1), we obtain:

$$P(|v - m| \geq m - p^* n) < \frac{\sigma^2}{(m - p^* n)^2}.$$  

Replacing $m$ and $\sigma$ by their values, as $np^* < np$ when $p^* < p$, by transitivity, we have the following inequality:

$$P(v \leq np^*) < \frac{pq}{n(p - p^*)^2}.$$  

Which leads to:

$$P(v \leq p^* n) \xrightarrow{n \to \infty} 0.$$  

So

$$P(s_\omega(\omega^*) > p^* n) \xrightarrow{n \to \infty} 1.$$  

This result gives in the limit a lower bound on the score of the true state of the world provided that the agents are equally R3-reliable. It is interesting because it ensures for some voting rules that the true state of the world belongs to the set of states returned by the rule. Consider for instance the following voting rules:

Definition 3 $(M$ and $Q_p)$ Let $E$ be a profile from a set of $n$ agents.

- The majority rule $M$ is defined as:
  $$M(E) = \{\omega \ s.t. \ s_\omega(\omega) > n/2\}.$$

- More generally, given $k \in [0,1]$, the $k$-quota rule $Q_k$ is defined as:
  $$Q_k(E) = \{\omega \ s.t. \ s_\omega(\omega) > kn\}.$$  

The majority rule $M$ corresponds to a specific quota rule, namely the 0.5-quota rule.

As a direct corollary to Proposition 1 we get:

Proposition 2 Let $E$ be a profile from a set of $n$ independent agents. If all agents have the same R3-reliability $p > k$, then the true state of the world belongs to the set of states returned by the $k$-quota rule in the limit, i.e.,

$$P(\omega^* \in Q_k(E)) \xrightarrow{n \to \infty} 1.$$  

Let us stress that this proposition only mentions the membership of the true state of the world in the result of the voting process, but it does not exclude that many other states can also appear in this result. Obviously, this is problematic from the truth tracking point of view. In particular, if each agent $i$ gives all the possible worlds ($\{K_i\} = W$), then for the corresponding profile $E$ we get all the possible worlds (for instance $Q_k(E) = W$ whatever $k$), which is not informative at all about the true state of the world.

The problem is due to the notion of R3-reliability that is not strong enough for the truth tracking purpose. Intuitively, asking the agents to give the true state of the world with a high probability is necessary but not sufficient since it does not prevent agents from giving (as models of their bases) too many states. Especially, an agent $i$ whose base is always a tautology ($\{K_i\} = W$), obviously carrying no information, is considered fully R3-reliable (i.e., her R3-reliability $p_i$ is equal to 1), which is unexpected. Thus a stronger notion of reliability is necessary. The following notion of R4-reliability is intended to this purpose:

- Let us note $q_{j,i}$ the probability that the world $\omega_j$ belongs to the set of models of the base of an agent $i$, i.e., $q_{j,i} = P(\omega_j \models K_i)$. If there is no ambiguity on the agent then we will simply note $q_j$ instead of $q_{j,i}$.

  - The incompetence $Q_i$ of an agent $i$ is the maximal probability that a world different from $\omega^*$ belongs to the set of models of her base, i.e., $Q_i = \max_{\omega \in \omega \setminus \{\omega^*\}} q_{j,i}$.

  - The competence of an agent is $c_i = 1 - Q_i$.

  - An agent is competent if $c_i > 0.5$. 

- The majority rule $M$ now is defined as:
  $$M(E) = \{\omega \ s.t. \ s_\omega(\omega) > n/2\}.$$  

- More generally, given $k \in [0,1]$, the $k$-quota rule $Q_k$ is defined as:
  $$Q_k(E) = \{\omega \ s.t. \ s_\omega(\omega) > kn\}.$$  

The majority rule $M$ corresponds to a specific quota rule, namely the 0.5-quota rule.
• An agent is R4-reliable if it is more R3-reliable than incompetent: $p_i > Q_i$. \hfill (R4)

Intuitively, while R3-reliability expresses the ability of an agent not to miss the true state of the world, the notion of competence deals with the quantity of uncertainty pervading her beliefs. Taken together, R3-reliability and competence are natural and important notions for characterizing the intuitive notion of “reliable agent” in the belief merging setting. While, in the specific case when $W$ consists only of two alternatives, an agent is competent if and only if she is R3-reliable, in the general case competence and R3-reliability are two different notions. Furthermore, it is easy to prove that the notion of R4-reliability extends the previous notions of reliability:

**Proposition 3**  
• When considering only profiles $E_c$ of complete bases, R4-reliability is equivalent to R2-reliability.

• When considering only profiles $E_c$ of complete bases and a set $W$ of interpretations containing only two elements $\{\omega, \omega^*\}$, R4-reliability, R3-reliability, R2-reliability and R1-reliability are equivalent.

With the notions of R4-reliability and competence, we can state the following Jury Theorem under Uncertainty:

**Theorem 3** Let $W = \{\omega_1, \ldots, \omega_{k-1}, \omega^*\}$ a set of possible worlds and let $E$ be a profile from a set of $n$ independent, homogenous and R4-reliable agents. Then $\forall i \in \{1, \ldots, k-1\},$

$$P(s_a(\omega^*) > s_a(\omega_i)) \longrightarrow_{n \rightarrow \infty} 1.$$ 

**Proof:** Let $(s_a(\omega_1), \ldots, s_a(\omega_{k-1}), s_a(\omega^*))$ be a vector of random variables where $s_a(\omega_i) = l (i \in \{1, \ldots, k-1\})$ (resp. $s_a(\omega^*) = l$) means that the score $s_a(\omega_i)$ (resp. $s_a(\omega^*)$) is equal to $l (l \in \{0 \ldots n\})$. As the set of agents is homogeneous, we have $q_{j,i} = q_{j,k}$ for every world $\omega_j$ and all agents $i, k$. We note $q_{j}$ this probability, i.e. $q_j = q_{j,i} = P(\omega_j \models K_i)$, for any agent $i$.

Each of the random variables $s_a(\omega_i) (i \in \{1, \ldots, k-1\})$ (resp. $s_a(\omega^*)$) follows a binomial distribution with parameters $n$ and $q_i$ (resp. $n$ and $p$). Subsequently, we have that $\forall j = 0 \ldots n$:

$$P(s_a(\omega_i) = j) = \binom{n}{j} q_i^j (1 - q_i)^{n-j}$$

and

$$P(s_a(\omega^*) = j) = \binom{n}{j} p^j (1 - p)^{n-j}.$$ 

The mean of each $s_a(\omega_i) (i \in \{1, \ldots, k-1\})$ is $nq_i$, its variance is $nq_i(1 - q_i)$, the mean of $s_a(\omega^*)$ is $np$ and its variance is $np(1 - p)$.

The (weak) law of large numbers applied to $s_a(\omega_i) (i \in \{1, \ldots, k-1\})$ and $s_a(\omega^*)$ gives that $\forall \epsilon > 0$:

$$P\left(\frac{s_a(\omega_i)}{n} - q_i \geq \epsilon\right) \longrightarrow_{n \rightarrow \infty} 0,$$ \hfill (2)

$$P\left(\frac{s_a(\omega^*)}{n} - p \geq \epsilon\right) \longrightarrow_{n \rightarrow \infty} 0.$$ \hfill (3)

Let $q = \max_{i=\{1, \ldots, k-1\}} q_i$ and $\epsilon_1 = \frac{q - p}{2}$. Since each agent is R4-reliable, we have that $q_i < p$ for each $i \in \{1, \ldots, k-1\}$, so $q < p$. As a consequence, we get that $\epsilon_1 > 0$. Taking advantage of inequations (2) and (3), one concludes that for each $i \in \{1, \ldots, k-1\}$,

$$P\left(\frac{s_a(\omega_i)}{n} - q_i \geq \epsilon_1\right) \longrightarrow_{n \rightarrow \infty} 0,$$

and

$$P\left(\frac{s_a(\omega^*)}{n} - p \geq \epsilon_1\right) \longrightarrow_{n \rightarrow \infty} 0.$$

It easily gives that:

$$P\left(\frac{s_a(\omega^*)}{n} > q_i + \epsilon_1\right) \longrightarrow_{n \rightarrow \infty} 0,$$ \hfill (4)

and

$$P\left(\frac{s_a(\omega^*)}{n} < p - \epsilon_1\right) \longrightarrow_{n \rightarrow \infty} 0.$$ \hfill (5)

The picture above explains the idea of the proof: when the weak law of large numbers can be used for a random variable, the values of this variable are closed to its mean with a high probability. Schematically, the probability that all the values of the variable are in a sphere with the mean as center and $\epsilon_1$ as radius tends to 1 in the limit. As a consequence, as $p > q$, the probability that the two spheres intersect tends to 0 in the limit.

Formally, suppose now that $\frac{s_a(\omega_i)}{n} \leq q_i + \epsilon_1$ and that $\frac{s_a(\omega^*)}{n} \geq p - \epsilon_1$, then, as $\forall i \in \{1 \ldots k-1\}, q_i + \epsilon_1 \leq p - \epsilon_1$, we get:

$$\frac{s_a(\omega_i)}{n} \leq q_i + \epsilon_1 \leq p - \epsilon_1 \leq \frac{s_a(\omega^*)}{n}.$$ 

The case when $\frac{s_a(\omega_i)}{n} > q_i + \epsilon_1$ or $\frac{s_a(\omega^*)}{n} < p - \epsilon_1$ may happen only if:

$$\frac{s_a(\omega_i)}{n} > q_i + \epsilon_1, \text{ or } \frac{s_a(\omega^*)}{n} < p - \epsilon_1.$$

And then we have:

$$P\left(\frac{s_a(\omega_i)}{n} > q_i + \epsilon_1\right) = P\left(\frac{s_a(\omega_i)}{n} > q_i + \epsilon_1 \text{ and } \frac{s_a(\omega^*)}{n} > \frac{s_a(\omega^*)}{n}\right) > P\left(\frac{s_a(\omega_i)}{n} > q_i + \epsilon_1 \text{ and } \frac{s_a(\omega^*)}{n} < p - \epsilon_1\right)$$

As $P\left(\frac{s_a(\omega_i)}{n} < p - \epsilon_1\right)$ and $\frac{s_a(\omega_i)}{n} > s_a(\omega^*)$ $\leq P\left(\frac{s_a(\omega^*)}{n} < p - \epsilon_1\right)$ and $P\left(\frac{s_a(\omega_i)}{n} > q_i + \epsilon_1 \text{ and } \frac{s_a(\omega^*)}{n} > \frac{s_a(\omega^*)}{n}\right)$. We get:

$$P\left(\frac{s_a(\omega_i)}{n} > q_i + \epsilon_1\right) \leq P\left(\frac{s_a(\omega^*)}{n} < p - \epsilon_1\right) + P\left(\frac{s_a(\omega^*)}{n} > q_i + \epsilon_1\right).$$

Finally, with assertions (4) and (5), we obtain:
Proposition 4

\[ \Delta \leq \Delta \]

or, equivalently

\[ P(\omega_i) \geq s_a(\omega^*) \quad \text{as } n \to \infty. \]

In this theorem, agents are not required to be R3-reliable or competent. Indeed, the conclusion holds as soon as the R3-reliability of each agent is greater than her incompetence.

This result is a generalization of Condorcet’s Jury Theorem to an uncertainty framework, where the agents can give a set of worlds instead of a single one. Interestingly, allowing the agents to vote for any number of worlds, and to choose as result the worlds with the greatest score is just approval voting [Brams and Fishburn, 1983]:

Definition 4 (Approval) Given a profile of bases E, the approval rule av is defined as:

\[ av(E) = \{ \omega \text{ s.t. } \forall \omega' \in W \quad s_a(\omega) \geq s_a(\omega') \}. \]

Thus, Theorem 3 shows that approval voting ensures to track the true world in the uncertain framework, just as plurality voting does in the usual framework.

5 Truth Tracking for Belief Merging

The ability of a merging operator \( \Delta \) to achieve the truth tracking issue can be modeled as a new postulate, called Truth Tracking postulate:

TT Let \( \omega^* \) be the true world. Let \( (E_n)_{n \in \mathbb{N}} \) be any sequence of profiles from a set of \( n \) independent, homogenous and R4-reliable agents. Then

\[ P(\Delta(E_n) = \{\omega^*\}) \quad \text{as } n \to \infty. \]

This postulate is satisfied by a merging operator when it allows to identify the true state of the world by listening sufficiently many homogeneous independent agents who are more R3-reliable than incompetent.

Let us now investigate the behaviour of some well-known belief merging operators with respect to this postulate. We first recall the definition of distance-based merging operators (see [Konečný and Pino Pérez, 2002] for details).

Definition 5 (distance-based merging operators) Let \( d \) be a pseudo-distance between worlds and \( f \) be an aggregation function. The merging operator \( \Delta_{d,f}(E) \) is defined by:

\[ \Delta_{d,f}(E) = \bigwedge_{i=1}^{\infty} \{ \mu_i \leq E \}, \]

where the pre-order \( \leq_E \) on \( W \) induced by \( E \) is defined by:

- \( \omega \leq_E \omega' \) if and only if \( d(\omega, E) \leq f(\omega', E) \), where
- \( d(\omega, E) = f_{K \in E}(d(\omega, K)) \), where
- \( d(\omega, K) = \min_{\omega' = K} d(\omega, \omega') \).

\( d_D \) denotes the drastic distance \( d_D(\omega, \omega') = 0 \) if \( \omega = \omega' \) and 1 otherwise, and \( d_H \) the Hamming distance.

The next result is a direct consequence of Theorem 3:

Proposition 4 \( \Delta_{d_D,\Sigma} \) satisfies TT.

Proposition 5 \( \Delta_{d_Q,\Sigma} \) does not satisfy TT.

Proof: Consider four possible worlds \( W = \{00, 01, 10, 11\} \), and suppose that the true state of the world is \( \omega^* = 11 \). Then consider the following probability distribution on the worlds, for all \( K_i \): \( P(\omega \mid K_i) = a > 0.5 \), \( P(\omega \mid K_i) = b, P(\omega \mid K_i) = c, P(\omega \mid K_i) = 0 \), with \( a + b + c = 1 \), \( b > 0 \), and \( c > 0 \). We consider also that the presence of any world in any base \( K_i \) is independent from the presence of any other world in the base (i.e., \( \forall \omega, \omega' \in W, P(\omega \mid K) = P(\omega \mid K_i) \)).

Since 10 never is a model of a belief base \( K_i \), the Hamming distance between 01 and a belief base never is greater than 1. Contrarily, the Hamming distance between \( \omega^* = 11 \) and the base whose set of models is \( \{00\} \) is 2. Since the probability of this base is \( b(1-a)(1-c) > 0 \), the probability that \( E \) contains such a base tends to 1 when \( n \) tends to infinity.

Hence \( d(01, E_n) < d(11, E_n) \) when \( n \) tends to infinity. This is enough to conclude that \( \lim_{n \to \infty} P(\Delta_{d_Q,\Sigma}(E_n) = \{11\}) \neq 1 \).

One easily concludes from the last two propositions that while IC merging operators aim at giving a synthesis of the input profile, not all of them are interesting to track the truth.

6 Some Empirical Results

The results about truth tracking we pointed out in the previous sections all concern the identification of the true state of the world \( \omega^* \) in the limit. None of them gives any information about truth tracking from the practical side, in the sense of a bound on the number of bases from which the identification is achieved with high probability.

In order to investigate this issue, we performed a number of experiments using \( \Delta_{d_D,\Sigma} \), that is an IC merging operator which satisfies TT and that is easy to implement. We investigated the convergence speed of truth tracking using \( \Delta_{d_D,\Sigma} \), depending on the agents R3-reliability \( p \) and incompetence \( Q = q \); for simplicity reasons, we made the assumption that all worlds \( \omega_i \) (different from the true world) have the same probability (i.e., \( \forall \omega_i \in W \setminus \{\omega^*\}, q_i = q \)). We considered sets of interpretations of various sizes (up to 2\(^{15}\)), we fixed the true state \( \omega^* \) as the world mapping each propositional variable to 0, and we generated profiles \( E \) from \( n \) homogeneous and independent agents, with R3-reliability \( p \) and incompetence \( q \) for different values of \( n \). For each value of \( n \), we computed 1000 profiles \( E \). For each \( E \), we computed \( \Delta_{d_D,\Sigma}(E) \) and check whether \( \{\omega^*\} \) holds. The proportion of the 1000 profiles for which \( \Delta_{d_D,\Sigma}(E) = \{\omega^*\} \) holds gives an estimate of the probability of success of truth tracking.

Figure 1 gives the probability that \( \{\Delta_{d_D,\Sigma}(E) = \{\omega^*\} \) given the number \( n \) of agents, when \( p = 0.3 \), and \( |W| =
2^7 worlds, for several values of q. Figure 2 is about similar experiments, but with p = 0.9.

Interestingly, we can observe on Figure 1 that even if p = 0.3 is rather low (here the agents under consideration are not R3-reliable), the convergence speed is high: to get \([\Delta_{D,\Sigma}(E)] = \{\omega^*\}\) with probability > 90%, 1250 agents are necessary for q = 0.25, 300 agents are necessary for q = 0.2 and only 60 agents are necessary for q = 0.1.

These results are very close to the ones reported on Figure 2 for p = 0.9. The corresponding values are 800 for q = 0.85, 230 for q = 0.8, and 40 for q = 0.6.

Empirically, it turns out that the "level of R4-reliability" of agents, i.e., the value p – q seems to have more impact on the convergence speed of truth tracking using \(\Delta_{D,\Sigma}\) than the fact that these values of p and q are rather high or rather low.

Figure 3 gives the probability of success of truth tracking given the number of propositional variables (hence the number of worlds) with p = 0.7 and q = 0.4.

As expected the complexity of discriminating the true state of the world increases with the number of possible states of the world. But, interestingly, the number of agents to be considered in order to achieve the truth tracking issue with high probability is not that huge compared to the number of interpretations. For instance, one can observe on Figure 3 that, when 10 variables are considered, less than 50 agents are enough to ensure that \([\Delta_{D,\Sigma}(E)] = \{\omega^*\}\) with probability greater than 90%, despite the fact that a single state has to be discriminated among 1024 ones and that the agents R3-reliability and competence are not so high.

7 Conclusion

In this work we have considered belief merging under the so-called "epistemic view", which amounts to evaluating the ability of belief merging operators to track the true state of the world. We proved a generalization of Condorcet’s Jury Theorem under uncertainty. We also defined a corresponding truth tracking postulate for belief merging and showed that some belief merging operators satisfies it. Finally we studied the convergence speed of \(\Delta_{D,\Sigma}\).

The problem of truth tracking has also been studied in the related framework of judgment aggregation [Bovens and Rabinowicz, 2006]. In [Pigozzi and Hartmann, 2007] the authors study the performances of the operator \(\Delta_{D,\Sigma}\) for this purpose, and show that it performs quite well (typically better than other judgement aggregation procedures) for mildly reliable agents.

The hypotheses considered in the Jury Theorem under Uncertainty may seem quite demanding. In most cases it is hard to ensure that agents are both reliable, independent and homogeneous. Nonetheless these hypotheses are exactly the same ones as those used in previous works (Condorcet and List-Goodin theorems). On the other hand, as for Condorcet’s Jury Theorem, one can expect generalizations to hold also for the Jury Theorem under Uncertainty; by relaxing for instance the independence assumption [Estlund, 1994] or the reliability one [Owen et al., 1989], such generalizations could be ob-
tained. Finally, the hypotheses we considered in the Jury Theorem under Uncertainty are enough for discarding some IC merging operators (especially $\Delta_{d,H}^d$,$G_{max}^H$) from those which could be suited to the truth tracking issue.

For future work, we plan to make some experiments in order to envision whether a jury theorem holds or not in relaxed cases. We also plan to study the truth tracking issue for other belief merging operators, such as the ones based on the Hamming distance. Our first experiments suggest that $\Delta_{d,H}^d$, $\Sigma$ also satisfies $TT$. Deriving a formal proof of it is an issue for further research.

References


