



Measuring Disagreement with Interpolants

Jandson S. Ribeiro^(✉), Viorica Sofronie-Stokkermans, and Matthias Thimm

University of Koblenz-Landau, Koblenz, Germany
{jandson,sofronie,thimm}@uni-koblenz.de

Abstract. We consider the problem of quantitatively assessing the conflict between knowledge bases in knowledge merging scenarios. Using the notion of Craig interpolation we define a series of disagreement measures and analyse their compliance with properties proposed in previous work by Potyka. We study basic complexity theoretic questions in that scenario and discuss the suitability of our approaches.

Keywords: Disagreement measure · Craig interpolation

1 Introduction

Inconsistencies arise easily whenever knowledge has to be aggregated from different sources [2,3,14]. Approaches to *belief merging* and *information fusion* address these issues by providing computational approaches for automatically resolving these issues in some sensible way. These fields bear a close relationship with the fields of *judgement* and *preference aggregation* [8,20] and also feature their own version of Arrow’s impossibility result [1], insofar that there cannot be any “rational” belief merging approach [7]. This calls for semi-automatic methods that take human background knowledge into account when knowledge has to be merged in order not to remove important pieces of information. In order to support the task of semi-automatic merging, we investigate approaches to *analyse* belief merging settings, i. e., approaches that can explain reasons for inconsistency and assess their severity. More specifically, we investigate *disagreement measures* [17], i. e., functions that take a knowledge base profile $P = \langle \mathcal{K}_1, \dots, \mathcal{K}_n \rangle$ as input and return a non-negative value that quantifies the severity of the disagreement between the different sources of information modelled by $\mathcal{K}_1, \dots, \mathcal{K}_n$. Disagreement measures are closely related to *inconsistency measures* [9,18,19], which themselves are functions that assess the severity of inconsistency in a single knowledge base. Disagreement and inconsistency measures can be used to help in debugging inconsistencies in semi-automated settings [4,5,10].

In this paper, we develop novel disagreement measures based on the concept of Craig interpolation [6]. Given two knowledge bases $\mathcal{K}_1, \mathcal{K}_2$ with $\mathcal{K}_1 \cup \mathcal{K}_2$ being inconsistent, an *interpolant* is a formula that concisely characterises an aspect of this inconsistency (we will provide formal definitions later). Thus, interpolants

play a similar role in analysing the disagreement between two knowledge bases as minimal inconsistent subsets do in analysing the inconsistency within a single knowledge base. As a matter of fact, minimal inconsistent subsets play a dominant role in many approaches to measuring inconsistency [12, 15, 21] and, therefore, it seems plausible to explore the use of interpolants in measuring disagreement. In order to do that, we consider the set of all interpolants (up to semantical equivalence) and define measures based on the size of that set and on the information content of the weakest/strongest interpolants (which are well-defined concepts due to the fact that the set of interpolants form a complete lattice). We show that our approaches provide meaningful results and comply with many of the rationality postulates introduced in [17]. We also undertake a small study of the computational complexity of several tasks relevant for our work, showing that (unsurprisingly) all of those are intractable. In summary, the contributions of this paper are as follows:

1. We make some general observations on interpolants in order to establish a framework suitable for measuring disagreement (Sect. 3).
2. We present novel disagreement measures based on interpolants (Sect. 4).
3. We investigate the compliance of these disagreement measures with rationality postulates from the literature (Sect. 5).
4. We investigate the computational complexity of several tasks pertaining to our disagreement measures (Sect. 6).

We introduce necessary preliminaries in Sect. 2 and conclude in Sect. 7.

We omit several proofs due to space restrictions. These can be found in an online appendix¹.

2 Preliminaries

Let At be some fixed propositional signature, i. e., a (possibly infinite) set of propositions, and let $\mathcal{L}(\text{At})$ be the corresponding propositional language constructed using the usual connectives \wedge (*conjunction*), \vee (*disjunction*), \rightarrow (*implication*), and \neg (*negation*). A literal is a proposition p or a negated proposition $\neg p$.

Definition 1. *A knowledge base \mathcal{K} is a finite set of formulas $\mathcal{K} \subseteq \mathcal{L}(\text{At})$. Let \mathbb{K} be the set of all knowledge bases.*

A clause is a disjunction of literals. A formula is in conjunctive normal form (CNF) if the formula is a conjunction of clauses. If Φ is a formula or a set of formulas we write $\text{At}(\Phi)$ to denote the set of propositions appearing in Φ . For a set $\Phi = \{\phi_1, \dots, \phi_n\}$ let $\bigwedge \Phi = \phi_1 \wedge \dots \wedge \phi_n$ and $\neg \Phi = \{\neg \phi \mid \phi \in \Phi\}$.

Semantics for a propositional language is given by *interpretations* where an *interpretation* ω on At is a function $\omega : \text{At} \rightarrow \{\text{true}, \text{false}\}$. Let $\Omega(\text{At})$ denote the set of all interpretations for At . An interpretation ω *satisfies* (or is a *model* of)

¹ <http://mthimm.de/misc/rst.dismes.proofs.pdf>.

an atom $a \in \text{At}$, denoted by $\omega \models a$, if and only if $\omega(a) = \text{true}$. The satisfaction relation \models is extended to formulas in the usual way. For $\Phi \subseteq \mathcal{L}(\text{At})$ we also define $\omega \models \Phi$ if and only if $\omega \models \phi$ for every $\phi \in \Phi$.

In the following, let Φ, Φ_1, Φ_2 be formulas or sets of formulas. Define the set of models $\text{Mod}(\Phi) = \{\omega \in \Omega(\text{At}) \mid \omega \models \Phi\}$. We write $\Phi_1 \models \Phi_2$ if $\text{Mod}(\Phi_1) \subseteq \text{Mod}(\Phi_2)$. Φ_1, Φ_2 are *equivalent*, denoted by $\Phi_1 \equiv \Phi_2$, if and only if $\text{Mod}(\Phi_1) = \text{Mod}(\Phi_2)$. Define the closure $\text{Cn}(\Phi)$ of a formula or set of formulas Φ via $\text{Cn}(\Phi) = \{\phi \mid \Phi \models \phi\}$. If $\text{Mod}(\Phi) = \emptyset$ we also write $\Phi \models \perp$ and say that Φ is *inconsistent* (or *unsatisfiable*).

3 Craig Interpolants

An important result in first-order logic is Craig's Interpolation Theorem [6].

Theorem 1 [6]. *Let ϕ, ψ be closed formulæ such that $\phi \models \psi$. Then there exists a closed formula I containing only predicate symbols, function symbols and constants occurring in both ϕ and ψ such that $\phi \models I$ and $I \models \psi$.*

Every formula I satisfying the property in Theorem 1 will be called an interpolant of ϕ and ψ . In the context of propositional logic, and of finite sets of propositional formulas, the concept of *interpolant* specializes as follows:

Definition 2. *Let Φ and Φ' be finite sets of propositional logic formulas. A formula ϕ is called an interpolant of Φ wrt. Φ' if*

1. $\Phi \models \phi$,
2. $\Phi' \cup \{\phi\} \models \perp$, and
3. $\text{At}(\phi) \subseteq \text{At}(\Phi) \cap \text{At}(\Phi')$

Consider, for instance, two sets $\Phi_1 = \{r \vee \neg p, \neg r \vee \neg q\}$ and $\Phi_2 = \{p, q\}$. The formula $p \rightarrow \neg q$ is an interpolant of Φ_1 wrt Φ_2 , as $\Phi_1 \models p \rightarrow \neg q$, $\Phi_2 \cup \{p \rightarrow \neg q\} \models \perp$ and $\text{At}(\{p \rightarrow \neg q\}) = \{p, q\} \subseteq \text{At}(\Phi_1) \cap \text{At}(\Phi_2)$.

Clearly, two finite sets Φ and Φ' of formulas in propositional logic have an interpolant if and only if $\Phi \cup \Phi'$ is unsatisfiable. Let $\mathbb{I}(\Phi, \Phi')$ denote the set of interpolants of Φ and Φ' .

Let φ be a formula and x a propositional symbol, we write $\varphi[x \mapsto \top]$ to denote that all occurrences of x in φ are replaced by \top . Analogously, $\varphi[x \mapsto \perp]$ means that all occurrences of x are replaced by \perp . For instance, for the formula $\varphi = p \vee \neg q$, we get $\varphi[p \mapsto \top] \equiv \top \vee \neg q$.

Definition 3. *Let ϕ be a propositional formula and $x \in \text{At}(\phi)$. We use the following notation:*

- $\exists x \phi := \phi[x \mapsto \perp] \vee \phi[x \mapsto \top]$;
- $\forall x \phi := \phi[x \mapsto \perp] \wedge \phi[x \mapsto \top]$.

Let Φ be a finite set of propositional formulæ, and let ϕ be the conjunction of all the formulæ in Φ . For every $x \in \text{At}(\phi)$, we use the following notation $\exists x \Phi := \exists x \phi$, $\forall x \Phi := \forall x \phi$.

Propositional logic allows uniform interpolation: For every formula ϕ and every set $\{y_1, \dots, y_m\} \subseteq \text{At}(\phi)$ there exists a propositional formula I_ϕ such that $\text{At}(I_\phi) \subseteq \{y_1, \dots, y_m\}$ with the property that for every formula ϕ' such that $\phi \wedge \phi' \models \perp$ and $\{y_1, \dots, y_m\} = \text{At}(\phi) \cap \text{At}(\phi')$, I_ϕ is an interpolant of ϕ and ϕ' . This follows from the following result (here formulated for finite sets of propositional formulæ).

Proposition 1. *Let Φ be a finite set of propositional formulas. Assume that $\text{At}(\Phi) = \{x_1, \dots, x_n, y_1, \dots, y_m\}$. Then the following hold:*

1. $\Phi \models \exists x_1 \dots \exists x_n \Phi$.
2. Let ψ be a propositional formula with $\text{At}(\psi) \subseteq \{y_1, \dots, y_m\}$ such that $\Phi \models \psi$. Then $\exists x_1 \dots \exists x_n \Phi \models \psi$.
3. $\forall x_1 \dots \forall x_n \Phi \models \Phi$ (i.e. $\forall x_1 \dots \forall x_n \Phi \models \phi_i$ for every formula $\phi_i \in \Phi$).
4. Let ψ be a formula with $\text{At}(\psi) \subseteq \{y_1, \dots, y_m\}$ such that $\psi \models \Phi$. Then $\psi \models \forall x_1 \dots \forall x_n \Phi$.

In the following, we define some auxiliary notions for interpolants and make some first observations regarding the structure of $\mathbb{I}(\Phi, \Phi')$.

Proposition 2. *Let Φ and Φ' be finite sets of formulas.*

1. If $\Phi \cup \Phi' \not\models \perp$ then $\mathbb{I}(\Phi, \Phi') = \emptyset$.
2. If $\Phi \models \perp$ and $\Phi' \models \perp$ then $\mathbb{I}(\Phi, \Phi') = \mathcal{L}(\text{At}(\Phi) \cap \text{At}(\Phi'))$.
3. If $\Phi \models \perp$ and $\Phi' \not\models \perp$ then $\mathbb{I}(\Phi, \Phi') = \mathcal{L}(\text{At}(\Phi) \cap \text{At}(\Phi')) \cap (\neg \text{Cn}(\Phi') \cup \{\phi \mid \phi \models \perp\})$.
4. If $\Phi \not\models \perp$ and $\Phi' \models \perp$ then $\mathbb{I}(\Phi, \Phi') = \mathcal{L}(\text{At}(\Phi) \cap \text{At}(\Phi')) \cap \text{Cn}(\Phi)$.

Proof. (1) Assume that $\mathbb{I}(\Phi, \Phi') \neq \emptyset$. Then there exists a formula ϕ such that $\Phi \models \phi$ and $\Phi' \cup \{\phi\} \models \perp$. Hence, $\Phi \cup \Phi' \models \perp$.

(2) If $\Phi \models \perp$ and $\Phi' \models \perp$ then (i) $\Phi \models \phi$ for all $\phi \in \mathcal{L}(\text{At}(\Phi) \cap \text{At}(\Phi'))$ and (ii) $\Phi' \cup \{\phi\} \models \perp$ for all $\phi \in \mathcal{L}(\text{At}(\Phi) \cap \text{At}(\Phi'))$.

(3) If $\Phi \models \perp$ then $\Phi \models \phi$ for all $\phi \in \mathcal{L}(\text{At}(\Phi) \cap \text{At}(\Phi'))$. If $\Phi' \not\models \perp$ then $\Phi' \cup \{\phi\} \models \perp$ for all ϕ such that (i) $\phi = \neg\phi'$, where $\phi' \in \text{Cn}(\Phi')$ or (ii) $\phi \models \perp$.

(4) If $\Phi \not\models \perp$ and $\Phi' \models \perp$ then a formula ϕ is an interpolant iff $\Phi \models \phi$ and $\phi \in \mathcal{L}(\text{At}(\Phi) \cap \text{At}(\Phi'))$.

As the notion of an interpolant is a syntactical one, the set $\mathbb{I}(\Phi, \Phi')$ also contains infinite syntactical variants for each interpolant (except for the case where $\Phi \cup \Phi'$ is consistent); in particular, if $\phi \in \mathbb{I}(\Phi, \Phi')$ then for every formula $\phi' \in \mathcal{L}(\text{At}(\Phi) \cap \text{At}(\Phi'))$, if $\phi \equiv \phi'$ then $\phi' \in \mathbb{I}(\Phi, \Phi')$. However, we will consider the following finite representation of $\mathbb{I}(\Phi, \Phi')$. For that, let $[\cdot] : \mathcal{L}(\text{At}) \rightarrow \Omega(\text{At})$ be a function that maps each formula to its *equivalence class* w.r.t. \equiv .

Definition 4. *Let Φ and Φ' be finite set of formulas. We denote by $\mathbb{S}\mathbb{I}(\Phi, \Phi')$ the set of equivalence classes of interpolants wrt. semantical equivalence, i. e.,*

$$\mathbb{S}\mathbb{I}(\Phi, \Phi') = \{[\phi] \mid \phi \in \mathbb{I}(\Phi, \Phi')\}$$

In the following, by abuse of notation we refer to $[\phi]$ also as an interpolant if ϕ is an interpolant.

Remark 1. If the sets Φ, Φ' of propositional formulæ are finite then the set $\text{At}(\Phi) \cap \text{At}(\Phi')$ is finite, so the set of all equivalence classes of formulæ over this set of atoms is finite and has at most $2^{2^{|\text{At}(\Phi) \cap \text{At}(\Phi')|}}$ elements, thus $\mathbb{S}\mathbb{I}(\Phi, \Phi')$ is finite.

It can be easily seen that the elements of $\mathbb{S}\mathbb{I}(\Phi, \Phi')$ form a lattice wrt. \models and the operations \wedge and \vee . Formally, we can make the following observations.

Proposition 3. *Let Φ and Φ' be finite set of formulas.*

1. If $[\phi], [\phi'] \in \mathbb{S}\mathbb{I}(\Phi, \Phi')$ then $[\phi \wedge \phi'] \in \mathbb{S}\mathbb{I}(\Phi, \Phi')$.
2. If $[\phi], [\phi'] \in \mathbb{S}\mathbb{I}(\Phi, \Phi')$ then $[\phi \vee \phi'] \in \mathbb{S}\mathbb{I}(\Phi, \Phi')$.
3. There is a uniquely defined $[\phi_w] \in \mathbb{S}\mathbb{I}(\Phi, \Phi')$ with $\phi' \models \phi_w$ for all $[\phi'] \in \mathbb{S}\mathbb{I}(\Phi, \Phi')$.
4. There is a uniquely defined $[\phi_s] \in \mathbb{S}\mathbb{I}(\Phi, \Phi')$ with $\phi_s \models \phi'$ for all $[\phi'] \in \mathbb{S}\mathbb{I}(\Phi, \Phi')$.

Conditions 1 and 2 of the proposition above are trivial. We illustrate the intuition of both conditions 3 and 4 with an example. Consider the knowledge bases $\mathcal{K}_1 = \{a, b\}$ and $\mathcal{K}_2 = \{a, \neg b\}$. We have that

$$\mathbb{S}\mathbb{I}(\mathcal{K}_1, \mathcal{K}_2) = \{[a \wedge b], [b], [a \rightarrow b], [a \leftrightarrow b]\}.$$

Note that every formula in $\mathbb{S}\mathbb{I}(\mathcal{K}_1, \mathcal{K}_2)$ implies $a \rightarrow b$ (the ϕ_w of condition 3). Similarly, the formula $a \wedge b$ implies all formulae in $\mathbb{S}\mathbb{I}(\mathcal{K}_1, \mathcal{K}_2)$, which makes $a \wedge b$ the formula $[\phi_s]$ of condition 4.

Remark 2. If the sets Φ, Φ' of propositional formulæ are finite then as $\mathbb{S}\mathbb{I}(\Phi, \Phi')$ is finite, $(\mathbb{S}\mathbb{I}(\Phi, \Phi'), \wedge, \vee)$ is a complete lattice.

We call $[\phi_w]$ the *weakest* interpolant and $[\phi_s]$ the *strongest* interpolant. It can be easily seen that

$$[\phi_w] = \left[\bigvee_{[\phi] \in \mathbb{S}\mathbb{I}(\Phi, \Phi')} \phi \right] \quad [\phi_s] = \left[\bigwedge_{[\phi] \in \mathbb{S}\mathbb{I}(\Phi, \Phi')} \phi \right]$$

We abbreviate the weakest interpolant of $\mathbb{S}\mathbb{I}(\Phi, \Phi')$ by $\text{Weakest}(\Phi, \Phi')$ and the strongest interpolant of $\mathbb{S}\mathbb{I}(\Phi, \Phi')$ by $\text{Strongest}(\Phi, \Phi')$. If $\mathbb{S}\mathbb{I}(\Phi, \Phi') = \emptyset$ both notions are undefined. We conclude this section with some further observations that will be useful in the remainder of the paper.

Proposition 4. *Let Φ, Φ' , and Φ'' be finite set of formulas.*

1. $\mathbb{I}(\Phi, \Phi') \subseteq \mathbb{I}(\Phi, \Phi' \cup \Phi'')$
2. $\mathbb{I}(\Phi, \Phi') \subseteq \mathbb{I}(\Phi \cup \Phi'', \Phi')$
3. $\mathbb{S}\mathbb{I}(\Phi, \Phi') \subseteq \mathbb{S}\mathbb{I}(\Phi, \Phi' \cup \Phi'')$

4. $\text{SI}(\Phi, \Phi') \subseteq \text{SI}(\Phi \cup \Phi'', \Phi')$
5. $\text{Weakest}(\Phi, \Phi') \models \text{Weakest}(\Phi, \Phi' \cup \Phi'')$
6. $\text{Weakest}(\Phi, \Phi') \models \text{Weakest}(\Phi \cup \Phi'', \Phi')$
7. $\text{Strongest}(\Phi, \Phi' \cup \Phi'') \models \text{Strongest}(\Phi, \Phi')$
8. $\text{Strongest}(\Phi \cup \Phi'', \Phi') \models \text{Strongest}(\Phi, \Phi')$

Proof. Considering item 1, for $\phi \in \mathbb{I}(\Phi, \Phi')$, from $\Phi' \cup \{\phi\} \models \perp$ it directly follows $\Phi' \cup \Phi'' \cup \{\phi\} \models \perp$. The other two conditions of Definition 2 remain valid as well. Item 2 is proven analogously and the remaining items follow directly from 2. \square

4 Disagreement Measures Based on Interpolation

We consider the scenario of measuring *disagreement* between multiple knowledge bases [17] using interpolation. For that, we denote a *knowledge base profile* by $P = \langle \mathcal{K}_1, \dots, \mathcal{K}_n \rangle$ with $\mathcal{K}_1, \dots, \mathcal{K}_n$ being knowledge bases. Furthermore, for P of this form and another knowledge base \mathcal{K} , we denote by $P \circ \mathcal{K}$ the concatenation of P with \mathcal{K} , i.e., $P \circ \mathcal{K} = \langle \mathcal{K}_1, \dots, \mathcal{K}_n, \mathcal{K} \rangle$. Furthermore, for $k \in \mathbb{N}$ define $P \circ^k \mathcal{K} = (P \circ^{k-1} \mathcal{K}) \circ \mathcal{K}$ and $P \circ^1 \mathcal{K} = P \circ \mathcal{K}$. Let \mathfrak{K} denote the set of all knowledge base profiles. Then Potyka [17] defines a disagreement measure as follows. Let $\mathbb{R}_{\geq 0}^{\infty}$ be the set of non-negative real values including ∞ .

Definition 5. A disagreement measure \mathcal{D} is a function $\mathcal{D} : \mathfrak{K} \rightarrow \mathbb{R}_{\geq 0}^{\infty}$ that satisfies

Consistency $\mathcal{D}(P) = 0$ iff $\bigcup P$ is consistent.

Symmetry $\mathcal{D}(\langle \mathcal{K}_1, \dots, \mathcal{K}_n \rangle) = \mathcal{D}(\langle \mathcal{K}_{\sigma(1)}, \dots, \mathcal{K}_{\sigma(n)} \rangle)$ for each permutation σ of $\{1, \dots, n\}$.

We also write $\mathcal{D}(\mathcal{K}_1, \dots, \mathcal{K}_n)$ for $\mathcal{D}(\langle \mathcal{K}_1, \dots, \mathcal{K}_n \rangle)$ to ease notation.

In the following, we define a series of disagreement measures that work on knowledge base profiles with exactly *two* elements. In order to generalise these measures to arbitrary knowledge base profiles, we consider the following constructions.

Definition 6. Let \mathcal{D} be a function $\mathcal{D} : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}_{\geq 0}^{\infty}$. Then the induced sum-measure $\mathcal{D}^{\Sigma} : \mathfrak{K} \rightarrow \mathbb{R}_{\geq 0}^{\infty}$ and max-measure $\mathcal{D}^{\max} : \mathfrak{K} \rightarrow \mathbb{R}_{\geq 0}^{\infty}$ are defined via

$$\mathcal{D}^{\Sigma}(P) = \sum_{\mathcal{K}, \mathcal{K}' \in P} \mathcal{D}(\mathcal{K}, \mathcal{K}')$$

$$\mathcal{D}^{\max}(P) = \max\{\mathcal{D}(\mathcal{K}, \mathcal{K}') \mid \mathcal{K}, \mathcal{K}' \in P\}$$

In order to obtain valid disagreement measures using these two constructions, we only need to require the *Consistency* property from the used two-place functions:

Consistency² $\mathcal{D}(\mathcal{K}_1, \mathcal{K}_2) = 0$ iff $\mathcal{K}_1 \cup \mathcal{K}_2$ is consistent.

Proposition 5. *Let \mathcal{D} be a function $\mathcal{D} : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}_{\geq 0}^{\infty}$ that satisfies *Consistency*² then both \mathcal{D}^{Σ} and \mathcal{D}^{\max} are disagreement measures.*

Proof. We first consider \mathcal{D}^{Σ} . If $\bigcup P$ is consistent then $\mathcal{K} \cup \mathcal{K}'$ is consistent as well for all $\mathcal{K}, \mathcal{K}' \in P$, implying $\mathcal{D}(\mathcal{K}, \mathcal{K}') = 0$. It follows $\mathcal{D}^{\Sigma}(P) = \mathcal{D}^{\max}(P) = 0$ showing *Consistency*. The property *Symmetry* is satisfied by construction.

Let us now define some concrete measure using interpolants.

Definition 7. *Let $\mathcal{K}_1, \mathcal{K}_2$ be finite and consistent set of formulas. Define the measure $\mathcal{D}_{\mathbb{S}\mathbb{I}} : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}_{\geq 0}^{\infty}$ via*

$$\mathcal{D}_{\mathbb{S}\mathbb{I}}(\mathcal{K}_1, \mathcal{K}_2) = |\mathbb{S}\mathbb{I}(\mathcal{K}_1, \mathcal{K}_2)|$$

Observe that $\mathcal{D}_{\mathbb{S}\mathbb{I}}$ satisfies *Consistency*², as $\mathbb{S}\mathbb{I}(\mathcal{K}, \mathcal{K}') = \emptyset$ if $\mathcal{K} \cup \mathcal{K}'$ is consistent. So $\mathcal{D}_{\mathbb{S}\mathbb{I}}^{\Sigma}$ and $\mathcal{D}_{\mathbb{S}\mathbb{I}}^{\max}$ are disagreement measures according to Proposition 5.

Example 1. Consider the three following knowledge bases: $\mathcal{K}_1 = \{a, b\}, \mathcal{K}_2 = \{b, c\}$ and $\mathcal{K}_3 = \{a, \neg b\}$. Going for their interpolants we get:

- $\mathbb{S}\mathbb{I}(\mathcal{K}_1, \mathcal{K}_2) = \mathbb{S}\mathbb{I}(\mathcal{K}_2, \mathcal{K}_1) = \emptyset$, as \mathcal{K}_1 is consistent with \mathcal{K}_2 ;
- $\mathbb{S}\mathbb{I}(\mathcal{K}_1, \mathcal{K}_3) = \{[b], [a \wedge b], [a \rightarrow b]\}$ and $\mathbb{S}\mathbb{I}(\mathcal{K}_3, \mathcal{K}_1) = \{[-b], [a \rightarrow \neg b], [a \wedge \neg b]\}$;
- $\mathbb{S}\mathbb{I}(\mathcal{K}_2, \mathcal{K}_3) = \{[b]\}$ and $\mathbb{S}\mathbb{I}(\mathcal{K}_3, \mathcal{K}_2) = \{[-b]\}$, as $\text{At}(\{\mathcal{K}_2, \mathcal{K}_3\}) = \{b\}$.

Therefore, $\mathcal{D}_{\mathbb{S}\mathbb{I}}^{\Sigma}(\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3) = 8$ and $\mathcal{D}_{\mathbb{S}\mathbb{I}}^{\max}(\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3) = 3$.

We consider another measure based on the information content of strongest (resp. weakest) interpolant. We use the following definition of an information measure, similar in spirit to the definition given in [11].

Definition 8. *An information measure J is a function $J : \mathcal{L}(\text{At}) \rightarrow \mathbb{R}_{\geq 0}^{\infty}$ that satisfies the following four properties:*

1. $J(\top) = 0$.
2. $J(\perp) = \infty$.
3. If $\phi \models \phi'$ then $J(\phi) \geq J(\phi')$.
4. If $\phi \not\models \perp$ then $J(\phi) < \infty$.

Here is a simple example of an information measures:

$$J_M(\phi) = \begin{cases} 0 & \text{if } \top \models \phi \\ \infty & \text{if } \phi \models \perp \\ 1/|\text{Mod}(\phi)| & \text{otherwise} \end{cases} \tag{1}$$

It is easy to verify that J_M is indeed an information measure according to Definition 8.

Then consider the following measure.

Definition 9. Let $\mathcal{K}_1, \mathcal{K}_2$ be finite and consistent set of formulas and J some information measure. Define the following measure $\mathcal{D}_J : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}_{\geq 0}^\infty$:

$$\mathcal{D}_J(\mathcal{K}_1, \mathcal{K}_2) = \begin{cases} 0 & \text{if } \mathcal{K}_1 \cup \mathcal{K}_2 \text{ is consistent} \\ \infty & \text{if } J(\text{Weakest}(\mathcal{K}_1, \mathcal{K}_2)) = 0 \\ \frac{J(\text{Strongest}(\mathcal{K}_1, \mathcal{K}_2))}{J(\text{Weakest}(\mathcal{K}_1, \mathcal{K}_2))} & \text{otherwise} \end{cases}$$

Again, observe that \mathcal{D}_J satisfies *Consistency*² (independently of J). So \mathcal{D}_J^\exists and \mathcal{D}_J^{\max} are disagreement measures according to Proposition 5.

Example 2. Consider the three knowledge bases from Example 1: $\mathcal{K}_1 = \{a, b\}, \mathcal{K}_2 = \{b, c\}$ and $\mathcal{K}_3 = \{a, \neg b\}$. In the table below, we show their strongest and weakest interpolants as well as their disagreement measure \mathcal{D}_{J_M} , where J_M is the information measure defined in (1) above.

	$(\mathcal{K}_1, \mathcal{K}_3)$	$(\mathcal{K}_3, \mathcal{K}_1)$	$(\mathcal{K}_2, \mathcal{K}_3),$	$(\mathcal{K}_3, \mathcal{K}_2)$	$(\mathcal{K}_1, \mathcal{K}_2)$	$(\mathcal{K}_1, \mathcal{K}_2)$
Weakest	$\neg a \vee b$	$\neg a \vee \neg b$	b	$\neg b$	-	-
Strongest	$a \wedge b$	$a \wedge \neg b$	b	$\neg b$	-	-
\mathcal{D}_{J_M}	3	3	1	1	0	0

As \mathcal{K}_1 is consistent with \mathcal{K}_2 they do not have any interpolant, and therefore their disagreement measure \mathcal{D}_{J_M} is zero. Note that each of the weakest interpolants between \mathcal{K}_1 and \mathcal{K}_3 have each 6 models, while their strongest interpolants have each 2 models.

5 Analysis

Potyka [17] proposes some desirable properties for disagreement measures, inspired by similar properties from inconsistency measurement [18]. Let \mathcal{D} be some disagreement measure. For $P = \langle \mathcal{K}_1, \dots, \mathcal{K}_n \rangle$, we say that $\mathcal{K}_i, i \in \{1, \dots, n\}$, is *involved in a conflict*, if there is $C \subseteq \{1, \dots, n\}$ such that $\bigcup_{j \in C} \mathcal{K}_j$ is consistent but $\mathcal{K}_i \cup \bigcup_{j \in C} \mathcal{K}_j$ is inconsistent.

Monotony (MO) $\mathcal{D}(\mathcal{K}_1, \dots, \mathcal{K}_n) \leq \mathcal{D}(\mathcal{K}_1 \cup \mathcal{K}', \dots, \mathcal{K}_n)$.

Dominance (DO) For formulas ϕ, ψ with $\phi \models \psi$ and $\phi \not\models \perp$, $\mathcal{D}(\mathcal{K}_1 \cup \{\phi\}, \dots, \mathcal{K}_n) \geq \mathcal{D}(\mathcal{K}_1 \cup \{\psi\}, \dots, \mathcal{K}_n)$.

Safe Formula Independence (SFI) For a formula ϕ with $\phi \not\models \perp$ and $\text{At}(\phi) \cap \text{At}(\bigcup \mathcal{K}_i) = \emptyset$, $\mathcal{D}(\mathcal{K}_1 \cup \{\phi\}, \dots, \mathcal{K}_n) = \mathcal{D}(\mathcal{K}_1, \dots, \mathcal{K}_n)$.

Adjunction Invariance (AI) For formulas ϕ, ψ , $\mathcal{D}(\mathcal{K}_1 \cup \{\phi, \psi\}, \dots, \mathcal{K}_n) = \mathcal{D}(\mathcal{K}_1 \cup \{\phi \wedge \psi\}, \dots, \mathcal{K}_n)$.

Tautology (TA) If \mathcal{K} is a knowledge base with $\top \models \mathcal{K}$ then $\mathcal{D}(P) \geq \mathcal{D}(P \circ \mathcal{K})$.

Contradiction (CO) If \mathcal{K} is inconsistent then $\mathcal{D}(P) \leq \mathcal{D}(P \circ \mathcal{K})$.

Majority (MAJ) If $\mathcal{K} \in P$ is consistent and involved in a conflict, then there is $k \in \mathbb{N}$ with $\mathcal{D}(P \circ^k \mathcal{K}) < \mathcal{D}(P)$.

Majority Agreement in the Limit (MAJL) If M is a maximal consistent subset of $\bigcup P$ then $\lim_{k \rightarrow \infty} \mathcal{D}(P \circ^k M) = 0$.

Table 1. Compliance of investigated disagreement measures with postulates.

	MO	DO	SFI	AI	TA	CO	MAJ	MAJL
\mathcal{D}_{SI}^Σ	✓	✗	✓	✓	✗	✓	✗	✗
\mathcal{D}_{SI}^{\max}	✓	✗	✓	✓	✓	✓	✗	✗
\mathcal{D}_J^Σ	✓	✗	✓	✓	✓	✓	✗	✗
\mathcal{D}_J^{\max}	✓	✗	✓	✓	✓	✓	✗	✗

We check compliance of the disagreement measures $\mathcal{D}_{SI}^\Sigma, \mathcal{D}_{SI}^{\max}, \mathcal{D}_J^\Sigma$ and \mathcal{D}_J^{\max} against the disagreement measures postulates mentioned above. For this, we introduce the concept of agreement between interpretations. We say that an interpretation w agrees with an interpretation w' modulo a set of propositional symbols X iff $w(p) = w'(p)$, for all $p \in X$. Note that the agreement relation is symmetric, that is, if w agrees with w' modulo X then w' agrees with w modulo X . We recall the Coincidence Lemma on propositional logic:

Lemma 1 (Coincidence Lemma). *If two interpretations w and w' agree with $\text{At}(\alpha)$, then $w \models \alpha$ iff $w' \models \alpha$.*

Proposition 6. *Let \mathcal{K} be knowledge base, and φ a consistent formula. If $\text{At}(\mathcal{K}) \cap \text{At}(\varphi) = \emptyset$, $\text{At}(\alpha) \subseteq \text{At}(\mathcal{K})$ and $\mathcal{K} \cup \{\varphi\} \models \alpha$ then $\mathcal{K} \models \alpha$.*

Proof. The case that \mathcal{K} is inconsistent is trivial. So we focus on the case that \mathcal{K} is consistent. Let then $w \in \text{Mod}(\mathcal{K})$ be an interpretation of \mathcal{K} , we will show that $w \in \text{Mod}(\alpha)$. As φ is consistent, it has at least one interpretation. Let $w' \in \text{Mod}(\varphi)$, and let w_0 be the following interpretation: $w_0(p) = w(p)$, if $p \in \text{At}(\mathcal{K})$; and $w_0(p) = w'(p)$, otherwise. Note that w_0 agrees with w modulo $\text{At}(\mathcal{K})$, which implies that $w_0 \models \mathcal{K}$. As $\text{At}(\mathcal{K}) \cap \text{At}(\varphi) = \emptyset$, we get that w_0 agrees with w' modulo $\text{At}(\varphi)$. This implies that $w_0 \models \varphi$. Therefore, as $w_0 \models \mathcal{K}$ and $w_0 \models \varphi$, we have that $w_0 \models \mathcal{K} \cup \{\varphi\}$. Thus, as $\mathcal{K} \cup \{\varphi\} \models \alpha$, we get that $w_0 \models \alpha$. As, by hypothesis, $\text{At}(\alpha) \subseteq \text{At}(\mathcal{K})$, we get that w agrees with w_0 modulo $\text{At}(\alpha)$. This implies, from Lemma 1, that $w \models \alpha$.

Theorem 2. *The compliance of the measures $\mathcal{D}_{SI}^\Sigma, \mathcal{D}_{SI}^{\max}, \mathcal{D}_J^\Sigma$, and \mathcal{D}_J^{\max} is as shown in Table 1.*

Note again that the proof of the above theorem can be found online².

Potyka [17] has defined disagreement measures from inconsistency measures. The idea is that the degree of disagreement in a knowledge profile would correspond to measuring the degree of inconsistencies between the knowledge bases in a knowledge profile. Potyka then discusses about some principles that disagreement measures should satisfy, and show that many of the disagreement measures induced from inconsistency measures fail to satisfy some of these principles. He

² http://mthimm.de/misc/rst_dismes-proofs.pdf.

then proposes the disagreement measure D_η , based on the η -inconsistency measure [13], to capture some of these principles. Following this line, we analyse here some desirable properties that disagreement measures on interpolants satisfy. We show that Potyka's D_η measure has some issues that are better handled by disagreement measures based on interpolants.

We start by addressing the Adjunction Invariance axiom. Breach of this axiom can lead to unintuitive behaviours. For instance, both $\{\alpha \wedge \neg\alpha\}$ and $\{\alpha, \neg\alpha\}$ should have the same disagreement value, since they present the same conflicts. Most of the disagreement measures induced from inconsistency measures analysed by Potyka, including his D_η measure, is not adjunctive invariant. On the other hand, as shown above, disagreement measures based on interpolants are adjunctive invariant. This is because interpolants are not syntax sensitive, instead they consider formulæ that share the same signature.

Potyka's disagreement measure D_η plateaus when each knowledge base in a knowledge profile conflicts with each other. As he himself criticizes, this scenario makes the measurement purely dependable on the size of the knowledge profile. Precisely, in that case $D_\eta(P) = (|P| - 1)/|P|$. In this scenario, a disagreement measure should still be able to distinguish an increase of conflicts, even if two knowledge profiles have the same size. We illustrate it in the following example.

Example 3. Let $P = \langle \mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3 \rangle$, and $P' = \langle \mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_4 \rangle$ be two knowledge profiles, where $\mathcal{K}_1 = \{a, b\}$, $\mathcal{K}_2 = \{\neg a, b\}$, $\mathcal{K}_3 = \{\neg b\}$ and $\mathcal{K}_4 = \{\neg b, \neg a\}$. Note that the four knowledge bases are consistent, but are inconsistent with each other. For the D_η disagreement measure, we would have $D_\eta(P) = D_\eta(P') = 2/3$.

Note that, against \mathcal{K}_1 and \mathcal{K}_2 , the knowledge base \mathcal{K}_4 presents more conflicts than \mathcal{K}_3 . Thus, though both P and P' have the same size, P' is more conflicting than P . A disagreement measure should be able to distinguish this difference of conflicts. The D_η measure, however, does not distinguish these conflicts, since when knowledge bases are pairwise inconsistent, the measure considers only the size of the knowledge profile, which is rather simplistic.

On the other hand, disagreement measures on interpolants are able to distinguish this tenuous difference of conflicts. To illustrate this, consider the $\mathcal{D}_{\text{SI}}^\Sigma$ measure. For the knowledge bases above we would get $\mathcal{D}_{\text{SI}}(\mathcal{K}_1, \mathcal{K}_2) = 4$, $\mathcal{D}_{\text{SI}}(\mathcal{K}_1, \mathcal{K}_3) = 1$, $\mathcal{D}_{\text{SI}}(\mathcal{K}_1, \mathcal{K}_4) = 4$, $\mathcal{D}_{\text{SI}}(\mathcal{K}_2, \mathcal{K}_3) = 1$ and $\mathcal{D}_{\text{SI}}(\mathcal{K}_2, \mathcal{K}_4) = 4$. Thus, $\mathcal{D}_{\text{SI}}(P)^\Sigma = (4 + 1 + 1) * 2 = 12$ and $\mathcal{D}_{\text{SI}}(P')^\Sigma = (4 + 4 + 4) * 2 = 24$.

This shows that disagreement measures on interpolants, such as $\mathcal{D}_{\text{SI}}^\Sigma$, present ways of distinguishing sensible conflicts between knowledge bases. Towards this end, a deeper investigation of which rational behaviours interpolants yield for disagreement measures is a path worth to explore.

6 Computational Complexity

To conclude our analysis, we now investigate the computational complexity of problems related to our novel measures.

Let us first consider some general observations on interpolants, which seem— to the best of our knowledge—not have been explicitly mentioned in the literature thus far.

Theorem 3. *Let Φ and Φ' be sets of formulas and let ϕ be a formula. Deciding whether ϕ is an interpolant of Φ wrt. Φ' is CONP -complete.*

Proof. For CONP -membership, consider the complement problem of deciding whether ϕ is not an interpolant of Φ wrt. Φ' and define the following non-deterministic algorithm. On instance (ϕ, Φ, Φ') we guess a triple (ω, ω', a) with interpretations ω, ω' , and an atom a . Then ϕ is not an interpolant of Φ wrt. Φ' if either

1. $\omega \models \Phi$ and $\omega \not\models \phi$,
2. $\omega' \models \Phi' \cup \{\phi\}$, or
3. $a \in \text{At}(\phi) \setminus (\text{At}(\Phi) \cap \text{At}(\Phi'))$.

Observe that each of these checks correspond to disproving one item of Definition 2 and that they can be done in polynomial time. It follows CONP -membership of the original problem.

For CONP -hardness, we reduce the problem TAUT to our problem. An instance to TAUT consists of a formula ψ and it has to be decided whether ψ is tautological. On input ψ we construct the instance (ϕ, Φ, Φ') for our problem (let a be a fresh atom not appearing in ψ) with

$$\phi = a \qquad \Phi = \{\psi \rightarrow a\} \qquad \Phi' = \{\neg a\}$$

It remains to show that ψ is tautological if and only if ϕ is an interpolant of Φ wrt. Φ' . Assume that ψ is tautological, then it follows $\Phi \models a$. Obviously, $\Phi' \cup \{a\}$ is inconsistent and ϕ only contains atoms of the shared vocabulary. It follows that ϕ is an interpolant of Φ wrt. Φ' . The other direction is analogous. \square

In [16] it is shown that, essentially, the size³ of an interpolant $\phi \in \mathbb{S}\mathbb{I}(\Phi, \Phi')$ is probably not bound polynomially in the size of both Φ and Φ' (very surprising results would follow in this case). This makes a characterisation of the complexity of various other computational problems hard. For example, a standard approach to decide whether some given formula ϕ would *not* be the strongest (weakest) interpolant, would be to guess another formula ϕ' (e. g., the actual strongest/weakest interpolant) and verify that ϕ' is an interpolant and $\phi' \models \phi$ ($\phi \models \phi'$). However, as ϕ' might be of exponential size this is not feasible to show membership in some class of the polynomial hierarchy or even PSPACE . We can therefore only provide a straightforward upper bound for the complexity of these and other problems of relevance to us.

Theorem 4. *Let Φ and Φ' be sets of formulas.*

1. *For a formula ϕ , the problem of deciding whether $\phi = \text{Strongest}(\Phi, \Phi')$ is in EXPSpace .*

³ More specifically, the smallest size of a $\phi' \in [\phi]$.

2. For a formula ϕ , the problem of deciding whether $\phi = \text{Weakest}(\Phi, \Phi')$ is in EXPSPACE.
3. The problem of determining $\text{Strongest}(\Phi, \Phi')$ is in FEXPSPACE.
4. The problem of determining $\text{Weakest}(\Phi, \Phi')$ is in FEXPSPACE.
5. The problem of counting $|\text{SII}(\Phi, \Phi')|$ is in FEXPSPACE.

Proof. We first show 5. For that, we enumerate (by reusing space) every subset $M \subseteq \Omega(\text{At})$ of interpretations (note that every interpretation is of polynomial size and there are exponentially many interpretations; therefore M is of exponential size). Each such set M characterises a formula ϕ_M and its equivalence class $[\phi_M]$ via $\text{Mod}(\phi_M) = M$. We can then easily verify whether ϕ_M is an interpolant of Φ wrt. Φ' , see Theorem 3 and the fact that $\text{CONP} \subseteq \text{EXPSPACE}$. In the positive case we add 1 to some counter. After enumerating all possible M , the counter is exactly the number $|\text{SII}(\Phi, \Phi')|$.

In order to prove 1–4, the above algorithm can easily be adapted. For 1., whenever we have verified a formula ϕ_M to be an interpolant, we can check whether $\phi_M \models \phi$ and $\phi \not\models \phi_M$. In that case, ϕ cannot be the strongest interpolant. Case 2 is analogous. Cases 3–4 can be realised by keeping track of the strongest (weakest) interpolant found so far and always comparing it newly discovered interpolants. \square

7 Summary and Conclusion

We investigated the problem of measuring disagreement in belief merging scenarios. For that, we made use of the concept of Craig interpolants and defined disagreement measures that consider the number of semantically equivalent interpolants between two knowledge bases and the information content in the strongest and weakest interpolants. We showed that our measures satisfy a number of desirable properties and we briefly discussed the computational complexity of related problems.

For future work, we will investigate the possibility of defining further measures based on interpolation and investigate their properties. Moreover, a deeper analysis of the differences of our measures with the measures proposed by Potyka [17] is needed. Precisely, our approach based on interpolants has a semantic perspective in assessing the culpability degree of the inconsistencies between two knowledge bases. We shall investigate what properties this semantic perspective brings upon these inconsistency measures. We will also explore algorithmic approaches to compute our measures and investigate applying our ideas to the area of inconsistency measurement [19].

Acknowledgements. The research reported here was partially supported by the Deutsche Forschungsgemeinschaft (grant DE1983/9-1).

References

1. Arrow, K.J.: *Social Choice and Individual Values*. Wiley, New York (1963)
2. Baral, C., Minker, J., Kraus, S.: Combining multiple knowledge bases. *IEEE Trans. Knowl. Data Eng.* **3**(2), 208–221 (1991)
3. Cholvy, L., Hunter, A.: Information fusion in logic: a brief overview. In: Gabbay, D.M., Kruse, R., Nonnengart, A., Ohlbach, H.J. (eds.) *ECSQARU/FAPR -1997*. LNCS, vol. 1244, pp. 86–95. Springer, Heidelberg (1997). <https://doi.org/10.1007/BFb0035614>
4. Corea, C., Delfmann, P.: Supporting business rule management with inconsistency analysis. In: *Proceedings of the Dissertation Award, Demonstration, and Industrial Track at BPM 2018 co-located with 16th International Conference on Business Process Management (BPM 2018)*, Sydney, Australia, September 9–14, 2018, pp. 141–147 (2018)
5. Corea, C., Thimm, M.: Towards inconsistency measurement in business rule bases. In: *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI 2020)*, June 2020
6. Craig, W.: Linear reasoning. A new form of the Herbrand-Gentzen theorem. *J. Symb. Log.* **22**(3), 250–268 (1957)
7. Díaz, A.M., Pérez, R.P.: Impossibility in belief merging. *Artif. Intell.* **251**, 1–34 (2017)
8. Everaere, P., Konieczny, S., Marquis, P.: Belief merging versus judgment aggregation. In: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)* (2015)
9. Grant, J.: Classifications for inconsistent theories. *Notre Dame J. Formal Logic* **19**(3), 435–444 (1978)
10. Grant, J., Hunter, A.: Measuring consistency gain and information loss in stepwise inconsistency resolution. In: Liu, W. (ed.) *ECSQARU 2011*. LNCS (LNAI), vol. 6717, pp. 362–373. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22152-1_31
11. Grant, J., Hunter, A.: Measuring the good and the bad in inconsistent information. In: *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI 2013)* (2013)
12. Hunter, A., Konieczny, S.: Measuring inconsistency through minimal inconsistent sets. In: Brewka, G., Lang, J. (eds.) *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning (KR 2008)*, Sydney, Australia, pp. 358–366. AAAI Press, Menlo Park, September 2008
13. Knight, K.: Measuring inconsistency. *J. Philos. Logic* **31**(1), 77–98 (2002)
14. Konieczny, S., Pino Pérez, R.: On the logic of merging. In: *Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning (KR 1998)*. Morgan Kaufmann (1998)
15. Kedian, M., Liu, W., Jin, Z.: A general framework for measuring inconsistency through minimal inconsistent sets. *Knowl. Inf. Syst.* **27**, 85–114 (2011)
16. Mundici, D.: Tautologies with a unique craig interpolant, uniform vs. nonuniform complexity. *Ann. Pure Appl. Logic* **27**(3), 265–273 (1984)
17. Potyka, N.: Measuring disagreement among knowledge bases. In: *Proceedings of the 12th International Conference on Scalable Uncertainty Management (SUM 2018)* (2018)
18. Thimm, M.: On the evaluation of inconsistency measures. In: Grant, J., Martinez, M.V. (eds.) *Measuring Inconsistency in Information*. Studies in Logic, vol. 73. College Publications, February 2018

19. Thimm, M.: Inconsistency measurement. In: Proceedings of the 13th International Conference on Scalable Uncertainty Management (SUM 2019), December 2019
20. Williamson, J.: Aggregating judgements by merging evidence. *J. Logic Comput.* **19**(3), 461–473 (2009)
21. Xiao, G., Ma, Y.: Inconsistency measurement based on variables in minimal unsatisfiable subsets. In: Proceedings of the 20th European Conference on Artificial Intelligence (ECAI 2012) (2012)