Proceedings, Fifteenth International Conference on
Principles of Knowledge Representation and Reasoning (KR 2016)

A MIS Partition Based Framework for Measuring Inconsistency

Said Jabbour[1], Yue Ma[2,3], Badran Raddaoui[4], Lakhdar Saïs[1] and Yakoub Salhi[1]
[1]CRIL CNRS UMR 8188, University of Artois, France
[2]LRI, Univ. Paris-Sud, CNRS, Université Paris-Saclay, France
[3] Key Laboratory of Symbolic Computation and Knowledge Engineering, Jilin University, China
[4]LIAS - ENSMA, University of Poitiers, France
{jabbour,sais,salhi}@cril.fr, yue.ma@lri.fr, badran.raddaoui@ensma.fr

Abstract

In this paper, we propose a general framework, both *parameterized* and *parameter-free*, for defining a family of fine-grained inconsistency measures for propositional knowledge bases. The parameterized approach allows to encompass several existing inconsistency measures as specific cases, by properly setting its parameter. And the parameter-free approach is defined to avoid the difficulty in choosing a suitable parameter in practice but still keeps a desired ranking for knowledge bases by their inconsistency degrees. The fine granularity of our framework is based on the notion of *MIS partition* that considers the inner structure of all the minimal inconsistent subsets of a knowledge base. Moreover, MinCostSAT-based encodings are provided, which enable the use of efficient SAT solvers for the computation of the proposed measures. We implement these algorithms and test them on some real-world datasets. The preliminary experimental results for a variety of inputs show that the proposed framework gives a wide range of possibilities for evaluating large knowledge bases.

Introduction

Reasoning about inconsistent knowledge bases (KBs) has been a long-standing challenge in the AI community. In recent years, measuring inconsistency has proved useful in diverse scenarios, including software specifications (Barragans-Martinez, Arias, and Vilas 2004), belief merging (Qi, Liu, and Bell 2005), news reports (Hunter 2006), integrity constraints (Grant and Hunter 2006; 2013), and multi-agents systems (Hunter, Parsons, and Wooldridge 2014; Jabbour, Ma, and Raddaoui 2014).

Inconsistencies are often unavoidable in real-world applications. To achieve a certain goal, an agent may need to cooperate with another agent, even in the presence of conflicts between them. In this case, the agent would prefer one that has the *least* disagreement with herself. For instance, suppose that an agent A with the support on two topics t_1 and t_2, written $A = \{t_1, t_2\}$, has to collaborate with one of the following two agents: $B_1 = \{\neg t_1, \neg t_2\}$, and $B_2 = \{\neg t_1, \neg t_1 \vee \neg t_2\}$. Clearly, A is in conflict with both, since B_1 and B_2 agree upon the topic $\neg t_1$ that contradicts with the topic t_1 of A. In this case, it is not desired to immediately conclude that both agents have a same conflict degree with A. Instead, some proper inconsistency measures are necessary.

Among many possible ways to define an inconsistency measure (Hunter 2006; Grant and Hunter 2008; Hunter and Konieczny 2010; Ma, Qi, and Hitzler 2011; Jabbour et al. 2014), minimal inconsistent subsets (MISes) are often used because a MIS forms a direct representation of an inconsistency core in a KB. For example, a classical measure I_{MI} (Hunter and Konieczny 2010) is defined as the number of MISes of a base $\mathcal{K}$, i.e., $I_{MI}(\mathcal{K}) = |MISes(\mathcal{K})|$, by which we have $I_{MI}(A \cup B_1) = I_{MI}(A \cup B_2) = 2$ for the previously mentioned example. Although useful for many scenarios (Hunter and Konieczny 2010; 2008; Mu, Liu, and Jin 2011), I_{MI} measure fails to recommend B_1 or B_2 for A. We claim that B_1 has more conflicts with A than B_2 because there are two independent contradictory topics between A and B_1 (i.e. $\{t_1, \neg t_1\}$, $\{t_2, \neg t_2\}$), each of which should be modified to make an agreement between A and B_1. However, the disagreement between A and B_2 (i.e. $\{t_1, \neg t_1\}$, $\{t_1, t_2, \neg t_1 \vee \neg t_2\}$) can be handled by revising only one topic, for instance, if A deletes t_1. The problem of no distinction between B_1 and B_2 for A is due to the fact that I_{MI} treats all MISes equally in terms of their contributions to the inconsistency degree. Recently, another measure, called I_{CC}, has been developed as a lower bound of all *standard inconsistency measures* (Jabbour, Ma, and Raddaoui 2014), by considering the most representative MISes. The I_{CC} metric can distinguish B_1 and B_2 for A because $I_{CC}(A \cup B_1) = 2$ and $I_{CC}(A \cup B_2) = 1$. However, if another agent is available, say $B_3 = \{\neg t_1, t_1 \rightarrow t_2\}$, we get $I_{CC}(A \cup B_3) = 1$. That is, the agent A can not distinguish between B_2 and B_3, which is again insufficient in the following aspect: A and B_2 have more groups of topics in conflict (i.e. $\{t_1, \neg t_1\}$, $\{t_1, t_2, \neg t_1 \vee \neg t_2\}$) than A and B_3 (i.e. $\{t_1, \neg t_1\}$). From the above discussion about I_{MI} and I_{CC}, a fine-grained analysis of the inner structures of MISes is clearly necessary.

Studies on inner structures of MISes have been performed in many different settings. For instance, a dependence relation among MISes has been identified by the fact that resolving some MISes allows automatic resolution of others (Benferhat, Dubois, and Prade 1995). In the context of ontology debugging, root or derived axioms for unsatisfiability have been distinguished (Kalyanpur et al. 2005). A graphical represen-

Copyright © 2016, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

tation of the relationships among justifications and axioms is provided (Bail et al. 2011). In this paper, we propose the notion of MIS partition that results in a general framework of inconsistency measures.

Our contribution can be summarized as follows:

- Based on the MIS partition, we define a new family of weighted inconsistency measures and show that these measures satisfy some rational properties, and capture several existing inconsistency measures.
- To overcome the difficulty in choosing a suitable parameter in practice, we further provide a parameter-free inconsistency measure that can preserve the desired ranking (see below for details) generated from the MIS partition.
- We present Minimum-Cost Satisfiability (MinCostSAT) based encodings so that we can benefit from cutting-edge SAT solvers for computing the proposed fine-grained inconsistency measures.
- An experimental study is conducted to show the relevance of the proposed measures for finely quantifying the conflict status of real-world KBs.

Preliminaries

A propositional language $\mathcal{L}$ is built over a finite set of propositional symbols $\mathcal{PS}$ using classical logical connectives $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow\}$. The symbol $\perp$ denotes contradiction. $a, b, c, \ldots$ represent atoms in $\mathcal{PS}$. A literal is an atom a or its negation $\neg a$. A clause C is a disjunction of literals: $C = a_1 \vee \ldots \vee a_n$. A formula α in *conjunctive normal form* (CNF) is a conjunction of clauses. Let $Var(\alpha)$ denotes the set of variables in α. An interpretation is a total function from $\mathcal{PS}$ to $\{true, false\}$. An interpretation $\mathcal{B}$ is a model of α iff $\mathcal{B}(\alpha) = true$. A KB $\mathcal{K}$ is a finite set of propositional formulas. For a set S, we denote by $|S|$ its cardinality, and by 2^S its power set. $\mathcal{K}$ is *inconsistent* if $\mathcal{K} \vdash \perp$, where $\vdash$ is the classical consequence relation. Minimal inconsistent subsets, defined below, are often used to analyze inconsistency in a KB.

Definition 1 *Let $\mathcal{K}$ be a KB and $M \subseteq \mathcal{K}$. M is a* Minimal Inconsistent Subset *(MIS) of $\mathcal{K}$ iff $M \vdash \perp$ and $\forall M' \subsetneq M$, $M' \nvdash \perp$. The set of all minimal inconsistent subsets of $\mathcal{K}$ is denoted $MISes(\mathcal{K})$.*

A formula $\alpha \in \mathcal{K}$ is called a *free formula* iff $\nexists M \in MISes(\mathcal{K})$ s.t. $\alpha \in M$. The class of free formulas of $\mathcal{K}$ is written $free(\mathcal{K}) = \mathcal{K} \setminus \bigcup MISes(\mathcal{K})$, and its complement is named *unfree formulas* set: $unfree(\mathcal{K}) = \mathcal{K} \setminus free(\mathcal{K})$.

An **inconsistency measure** I is a function that maps a KB to a non-negative real number such that higher value indicates larger conflict.

Several desired properties have been defined to characterize inconsistency measures (Hunter and Konieczny 2010; Jabbour, Ma, and Raddaoui 2014; Besnard 2014). In this paper, given an arbitrary inconsistency measure I we focus on the following properties:

- *Consistency:* $I(\mathcal{K}) = 0$ iff $\mathcal{K}$ is consistent.
- *Monotonicity:* if $\mathcal{K} \subseteq \mathcal{K}'$, then $I(\mathcal{K}) \leq I(\mathcal{K}')$.
- *Independence:* $I(\mathcal{K} \cup \{\alpha\}) = I(\mathcal{K})$ if $\alpha \in free(\mathcal{K} \cup \{\alpha\})$.
- *MinInc:* $I(M) = 1$ if $M \in MISes(\mathcal{K})$.
- *Ind-decomposability:* $I(\mathcal{K}_1 \cup \ldots \cup \mathcal{K}_n) = \Sigma_{i=1}^{n} I(\mathcal{K}_i)$ if $MISes(\mathcal{K}_1 \cup \ldots \cup \mathcal{K}_n) = MISes(\mathcal{K}_1) \uplus \ldots \uplus MISes(\mathcal{K}_n)$, where $\uplus$ is the multi-set union over sets, and $unfree(\mathcal{K}_i) \cap unfree(\mathcal{K}_j) = \emptyset$ for $1 \leq i \neq j \leq n$.

The first four properties seem natural for an inconsistency measure. The idea behind the Ind-decomposability (Jabbour, Ma, and Raddaoui 2014) is that the inconsistency degrees of several KBs should be additive if these bases are disjoint and have disjoint MISes. Notice that this revised property is introduced to overcome the limitations of the Decomposability property (Hunter and Konieczny 2010). Another property named Dominance says that $I(\mathcal{K} \cup \{\alpha\}) \geq I(\mathcal{K} \cup \{\beta\})$ if $\alpha \vdash \beta$, which is however criticized due to its unsuitability for characterizing inconsistency measures based on minimal inconsistent subsets (Mu et al. 2011; Besnard 2014). Therefore, in the rest of this paper, we only consider the five postulates given above which are widely accepted by the AI community.

Definition 2 *An inconsistency measure is called a* standard measure *if it satisfies the Consistency, Monotonicity, Independence, MinInc, and Ind-decomposability properties.*

It is easy to check that I_{MI} is a standard measure.

MIS-based Measure I_{CC} Revisited

In (Jabbour, Ma, and Raddaoui 2014), the authors proposed a new inconsistency measure, denoted I_{CC}, based on a subtle analysis of the dependencies among the formulas of the KB. For example, in the I_{CC} measure, overlap among MISes are taken into account. Indeed, to characterize the inner structure of MISes, each KB $\mathcal{K}$ is associated with an hypergraph $G_{\mathcal{K}}$ whose vertices are associated with the formulas of $\mathcal{K}$ and edges with the MISes of $\mathcal{K}$. This hypergraph based representation gives a better insight of the correlations among MISes so that it becomes useful for analyzing inconsistencies. For example, the notion of *strong-partition* of a KB can be considered in the light of the connected components of $G_{\mathcal{K}}$, based on which the I_{CC} measure can be defined.

Definition 3 *(Jabbour, Ma, and Raddaoui 2014) Let $\mathcal{K}$ be a KB and $R \subseteq \mathcal{K}$. A* strong-partition *of $\mathcal{K}$ is a pair $\langle\{\mathcal{K}_1, \ldots, \mathcal{K}_n\}, R\rangle$ such that:*

(1) *$\mathcal{K}_i \subseteq \mathcal{K}$ and $\mathcal{K}_i \vdash \perp$, $\forall i$ $(1 \leq i \leq n)$;*
(2) *$\{\mathcal{K}_1, \ldots, \mathcal{K}_n, R\}$ is a partition of $\mathcal{K}$;*
(3) *$MISes(\mathcal{K}_1 \cup \ldots \cup \mathcal{K}_n) = \biguplus_{i=1}^{n} MISes(\mathcal{K}_i)$.*

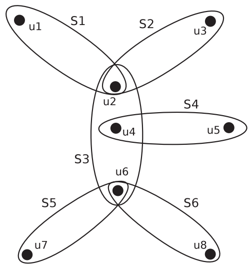
The I_{CC} measure is defined as $I_{CC}(\mathcal{K}) = m$ if there is a strong-partition $\langle D, R\rangle$ where $|D| = m$, and there is no strong-partition $\langle D', R'\rangle$ such that $|D'| > m$.

That is, when removing the set R from $\mathcal{K}$, I_{CC} corresponds exactly to the number of connected components of the hypergraph representing $\mathcal{K} \setminus R$.

One thing to note is that the measure I_{CC} is defined to serve as a lower bound of all standard measures.

Proposition 1 ((Jabbour, Ma, and Raddaoui 2014)) *For any KB $\mathcal{K}$ and any standard measure I, $I_{CC}(\mathcal{K}) \leq I(\mathcal{K})$.*

Example 1 *Consider the KB* $\mathcal{K} = \{a, \neg a, a \wedge b, (a \vee c) \wedge d, \neg d, \neg c \wedge e, \neg e, \neg e \wedge f\}$ *with its formulas named* $u_1, u_2, \ldots, u_8$*, respectively.* $\mathcal{K}$ *has six MISes as depicted in the hypergraph below. It follows that* $I_{\mathcal{CC}}(\mathcal{K}) = 2$*, which means that* $MISes(\mathcal{K})$ *are highly correlated, that is, the hypergraph can not be decomposed into more connected components even if some formulas are removed.*



In the following, we provide some additional properties of $I_{\mathcal{CC}}$ using the *closed set packing* problem defined in (Jabbour et al. 2015).

We first review the classical *set packing* problem defined as follows.

Definition 4 *Let* U *be a universe and* S *a family of subsets of* U*. A* set packing *is a subset* $P \subseteq S$ *such that,* $\forall S_i, S_j \in P$ *with* $S_i \neq S_j$*,* $S_i \cap S_j = \emptyset$*.*

The *closed set packing* problem is defined to circumscribe a special type of set packing.

Definition 5 *Let* U *be a universe and* S *a family of subsets of* U*. We define the function* $f_S : 2^S \rightarrow 2^S$ *as* $f_S(P) = \{S_i \in S \mid S_i \subseteq \cup_{S' \in P} S'\}$*. Then, a set packing* $P \subseteq S$ *is called a* closed set packing *(CSP) if* P *is a fixed point of the function* f_S*, i.e.,* $f_S(P) = P$*.*

That is, a CSP P of S satisfies the following condition: the union of the selected subsets does not contain any unselected subsets of S, i.e., $\forall S_i \in S \setminus P$, $S_i \not\subseteq \bigcup_{S' \in P} S'$.

Example 2 (Example 1 contd.) *By taking the universe* $U = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8\}$ *and* $S = MISes(\mathcal{K})$*, various CSPs can be constructed. For instance,* $\{S_2, S_4\}$ *is one, but not* $\{S_2, S_4, S_6\}$ *since* $S_3 \subseteq S_2 \cup S_4 \cup S_6$*. Indeed,* $\{S_2, S_4\}$ *is a CSP of maximal cardinality.*

The *maximum set packing (resp. closed set packing) problem*, written MSP (resp. MCSP), are the related optimization problems, defined as finding a set packing (resp. closed set packing) of S with the maximum size for a collection of subsets S over a universe U.

Given U and S, we write $\sigma_{MSP}(U, S)$ (resp. $\sigma_{MCSP}(U, S)$) for the cardinality of the optimal solution of the corresponding MSP (resp. MCSP) problem. Then, we have the following relationship between MSP and MCSP:

Proposition 2 *Given an inconsistent KB* $\mathcal{K}$*, it holds that:*

$$\sigma_{MCSP}(\mathcal{K}, MISes(\mathcal{K})) = \max_{R \subset \mathcal{K}} \{\sigma_{MSP}(\mathcal{K} \setminus R, MISes(\mathcal{K} \setminus R))\}.$$

Proof: (Sketch) The $\leq$ direction can be seen from the fact that for any R, a set packing w.r.t. $U = \mathcal{K} \setminus R$ and $S = MISes(\mathcal{K} \setminus R)$ is closed under the function f_S. The $\geq$ direction holds because a closed set packing $\mathcal{T}$ for $U = \mathcal{K}$ and $S = MISes(\mathcal{K})$ is indeed a set packing for $U = \mathcal{K} \setminus R$ and $S = MISes(\mathcal{K} \setminus R)$ with $R = \mathcal{K} \setminus \bigcup_{Q \subseteq \mathcal{T}} Q$. □

For Example 1, if we take $R = \{u_1, u_4, u_7\}$, we will get the maximal set packing of $\mathcal{K} \setminus R = \{s_2, s_6\}$ which is indeed a maximal closed set packing of $\mathcal{K}$. This proposition says that the optimal closed set packing corresponds to the maximal set packing with some sub-bases removed.

Using Proposition 2 and Definition 3, the $I_{\mathcal{CC}}$ measure can be nicely characterized by the closed set packing problem as stated by the following proposition.

Corollary 3 *For any KB* $\mathcal{K}$*, we have:*

$$I_{\mathcal{CC}}(\mathcal{K}) = \sigma_{MCSP}(\mathcal{K}, MISes(\mathcal{K})).$$

Proof: (Sketch) By Definition 3, $I_{\mathcal{CC}}$ value corresponds to the size of the largest subset of $MISes(\mathcal{K})$ having pairwise disjoints $MISes$ and closed by union. This corresponds to the solutions of the defined MCSP(U, S). □

The last result shows that $I_{\mathcal{CC}}$ value can be computed by leveraging solutions to the closed set packing problem. The detailed algorithm is given in the section dedicated to the computation issues.

Towards Weight-Based Standard Measures

As mentioned earlier, the $I_{\mathcal{CC}}$ measure defines a lower bound for standard measures. Unfortunately, the lower bound considers only a subset of MISes forming a closed set packing of MISes (cf. Corollary 3). That is, $I_{\mathcal{CC}}$ does not take into account the contribution of each MIS to the whole inconsistency. A key step for designing a more accurate inconsistency metric is to analyse the contribution of each MIS to the inconsistency of a given KB.

One of such analyses can be done via a weighted assessment of the relevance of each MIS by taking into account the overall structure of the whole set of MISes. To illustrate this point, let us consider the KB $\mathcal{K} = \{a, \neg a, a \vee b, \neg b\}$ with two MISes $M_1 = \{a, \neg a\}$, and $M_2 = \{\neg a, a \vee b, \neg b\}$. Since $M_1 \cap M_2 \neq \emptyset$, a "relevant" standard inconsistency measure I should satisfy $I(\mathcal{K}) < I(M_1) + I(M_2) = 2$. On the other hand, $I(\mathcal{K})$ should be greater than the lower bound $I_{\mathcal{CC}}(\mathcal{K}) = 1$. Consequently, we need to have[1] $1 < I(\mathcal{K}) < 2$. To satisfy such a condition, a deeper analysis of these two MISes by considering for example their overlap is necessary. This can be expressed using a weighted measure such as $I(\mathcal{K}) = w_1 \times I(M_1) + w_2 \times I(M_2)$ where $w_1, w_2 \in [0, 1]$ are positive real values representing the relevance of M_1 and M_2, respectively.

In this section, we develop a new weight-based standard inconsistency measure that captures the correlation between MISes. More precisely, we are interested in inconsistency metrics that assign higher (resp. lower) conflict degrees to KBs with sparser (resp. denser) MISes under the notion of hypergraph. The intuition is that a sparse MISes hypergraph

[1]Note that the Ind-decomposability cannot be applied to $\mathcal{K}$ because its MISes are inner connected. Hence, for a standard measure I, it is not required that $I(\mathcal{K}) = 2$.

indicates that the conflicts spread over the whole KB with various independent sources of inconsistency. To resolve such sort of inconsistency, all these conflicts need to be handled separately. To this end, our proposed approach exploits the closed set packing problem to characterize the correlation among different MISes.

Before defining our inconsistency measures, we need to introduce a *MIS partition* that partitions the set of MISes of a KB into clusters of closed set packing, called *c-partition*.

Definition 6 *Let $\mathcal{K}$ be a KB, $U = \mathcal{K}$ and $S = MISes(\mathcal{K})$. $\mathcal{P} = \{p_1, \ldots, p_n\}$ is called a* c-partition *of $\mathcal{K}$ if $MISes(\mathcal{K}) = \biguplus_{1 \leq i \leq n} p_i$, where each p_i is a closed set packing for U and S.*

In the sequel, let $\mathcal{P}_{MISes}(\mathcal{K})$ denote the set of c-partitions of $\mathcal{K}$. Obviously, $\mathcal{P}_{MISes}(\mathcal{K}) \neq \emptyset$ if $\mathcal{K} \vdash \bot$. Indeed, we can build a c-partition where each p_i contains exactly one MIS.

Now, we will associate with each c-partition $\mathcal{P} = \{p_1, \ldots, p_n\}$ the following numeric vector $V(\mathcal{P}) = \langle |\pi(p_1)|, \ldots, |\pi(p_n)| \rangle$ where π is a permutation of $\mathcal{P}$ such that $|\pi(p_1)| \geq \ldots \geq |\pi(p_n)|$. Clearly, each c-partition possesses a unique ordered numeric vector.

In order to compare two c-partitions of a given KB, it will be convenient to assume a lexicographic ordering relation $\preceq$ over their associated numeric vectors:

Definition 7 *Let $u, v \in \mathbb{N}^n \times \mathbb{N}^m$ be two vectors. Suppose that $u = \langle u_1, \ldots, u_n \rangle$ and $v = \langle v_1, \ldots, v_m \rangle$. Then, u is* lexicographically less *than v, denoted by $u \preceq v$, iff $u = v$, or there exists $k \leq \min(n, m)$ s.t. $u_k < v_k$ and $u_i = v_i$ for each $i < k$. Furthermore, $u \prec v$ iff $u \preceq v$ and $u \neq v$.*

Definition 8 *Let $\mathcal{K}$ be a KB and $\mathcal{P} = \{p_1, \ldots, p_n\}$, $\mathcal{P}' = \{p'_1, \ldots, p'_m\}$ two c-partitions of $\mathcal{K}$. Then, $\mathcal{P} \preceq_{pr} \mathcal{P}'$ iff $V(\mathcal{P}) \preceq V(\mathcal{P}')$.*

Note that the relation $\preceq_{pr}$ is a total preorder on $\mathcal{P}_{MISes}(\mathcal{K})$. Indeed, $\preceq_{pr}$ is reflexive and transitive. Additionally, no pair of c-partitions is incomparable.

Definition 9 *Let $\mathcal{K}$ be a KB and $\mathcal{P}$ be a c-partition of $\mathcal{K}$. Then, $\mathcal{P}$ is called a* maximal c-partition *if and only if there exists no c-partition $\mathcal{P}'$ such that $\mathcal{P} \preceq_{pr} \mathcal{P}'$.*

In the following, based on $\preceq_{pr}$, we can associate with each KB a vector of positive integers, called *conflict vector*.

Definition 10 *Let $\mathcal{K}$ be a KB. The* conflict vector *of $\mathcal{K}$, written $V(\mathcal{K})$, is defined as $V(\mathcal{K}) = V(\mathcal{P})$ where $\mathcal{P}$ is a maximal c-partition of $\mathcal{K}$.*

Example 3 (Example 1 contd.) *Let $\mathcal{K} = \{a, \neg a, a \wedge b, (a \vee c) \wedge d, \neg d, \neg c \wedge e, \neg e, \neg e \wedge f\}$. Then, $\mathcal{K}$ has 8 maximal c-partitions:*

$$\begin{array}{ll}
\mathcal{P}_1 = \{\{S_1, S_4\}, \{S_2, S_5\}, S_3, S_6\} & \mathcal{P}_2 = \{\{S_1, S_4\}, \{S_2, S_6\}, S_3, S_5\} \\
\mathcal{P}_3 = \{\{S_2, S_4\}, \{S_1, S_5\}, S_3, S_6\} & \mathcal{P}_4 = \{\{S_2, S_4\}, \{S_1, S_6\}, S_3, S_5\} \\
\mathcal{P}_5 = \{\{S_1, S_5\}, \{S_2, S_4\}, S_3, S_6\} & \mathcal{P}_6 = \{\{S_1, S_5\}, \{S_2, S_6\}, S_3, S_4\} \\
\mathcal{P}_7 = \{\{S_2, S_5\}, \{S_1, S_4\}, S_3, S_6\} & \mathcal{P}_8 = \{\{S_2, S_5\}, \{S_1, S_6\}, S_3, S_4\}
\end{array}$$

It follows that $V(\mathcal{K}) = \langle 2, 2, 1, 1 \rangle$.

Note that all the maximal c-partitions possess the same numeric vector. So the conflict vector of $\mathcal{K}$ is unique.

For two KBs with V($\mathcal{K}$) $= \langle |p_1|, \ldots, |p_n| \rangle$ and V($\mathcal{K}'$) $= \langle |p'_1|, \ldots, |p'_m| \rangle$, if $|p_1| = |p'_1|$, then $I_{\mathcal{CC}}$ associates the same inconsistency value to both $\mathcal{K}$ and $\mathcal{K}'$, since $I_{\mathcal{CC}}(\mathcal{K}) = |p_1|$ and $I_{\mathcal{CC}}(\mathcal{K}') = |p'_1|$ by Corollary 3. In contrast, the ordering defined above allows us to compare two KBs according to the remaining elements of the conflict vectors.

To quantify the conflict degree of a KB, we will use the set of c-partitions with an associated weighting scheme. This will allows us to better quantify the contribution of the different MISes to the inconsistency of the KB.

Definition 11 *Let $\mathcal{K}$ be a KB and $\mathcal{P} = \{p_1, \ldots, p_n\} \in \mathcal{P}_{MISes}(\mathcal{K})$. We define the* weighting function *$\mathcal{W}$ of $\mathcal{P}$ as:*

$$\mathcal{W}(\mathcal{P}) = \sum_{p_i \in \mathcal{P}} |p_i| \times w_i,$$

where $\{w_n\}_{n=1}^{+\infty}$ is a decreasing positive sequence with $w_1 = 1$.

That is, the MISes belonging to the same p_i are equally weighted and the weight associated to each c-partition is the sum of the individual weight associated to each MIS.

Now, we are ready to define our new class of inconsistency metrics induced by $\mathcal{W}$ as follows:

Definition 12 *Let $\mathcal{K}$ be a KB. The* weighted inconsistency measure *of $\mathcal{K}$, w.r.t. a weighting function $\mathcal{W}$, is defined as follows:*

$$I_{\mathcal{W}}(\mathcal{K}) = \max\{\mathcal{W}(\mathcal{P}) \mid \mathcal{P} \in \mathcal{P}_{MISes}(\mathcal{K})\}.$$

As we can remark, another possibility in defining $I_{\mathcal{W}}$ would be $\mathcal{W}(\mathcal{P}')$, where $\mathcal{P}'$ is a maximal c-partition of $\mathcal{K}$. However, these two ways yield different definitions in general. For instance, if $V(\mathcal{K}) = V(\mathcal{P}) = \langle 4, 1, 1, 1 \rangle$ for a maximal c-partition $\mathcal{P}$ and there exists another c-partition $\mathcal{P}'$ with $V(\mathcal{P}') = \langle 3, 3, 1 \rangle$, we have $I_{\mathcal{W}}(\mathcal{P}) - I_{\mathcal{W}}(\mathcal{P}') = 1 * w_1 - 2 * w_2 + w_4$. Now let us consider the weighting sequence $w_1 = 1, w_2 = 0.9$ and $w_4 = 0.1$, then $I_{\mathcal{W}}(\mathcal{P}) - I_{\mathcal{W}}(\mathcal{P}') = 1 - 1.8 + 0.1 = -0.7 < 0$. This means that the maximum value of $I_{\mathcal{W}}$ is not reached for the maximal c-partition. Moreover, as shown in Proposition 6 and Theorem 7, the measure $I_{\mathcal{W}}$, as defined in Definition 12, is general enough to encompass particular measures ($I_{\mathcal{CC}}, I_{MI}$), which nevertheless cannot be always guaranteed by $\mathcal{W}(\mathcal{P})$ for maximal c-partitions $\mathcal{P}$.

Given an inconsistent KB $\mathcal{K}$, It is important to note that the measure $I_{\mathcal{W}}$ reaches the maximum value when $MISes(\mathcal{K})$ itself forms a CSP. Indeed, in this case $MISes(\mathcal{K})$ has a singleton maximal c-partition $\mathcal{P}$ so that the conflict vector $V(\mathcal{K}) = \langle |MISes(\mathcal{K})| \rangle$. This property is meaningful as it indicates the case where the sources of conflicts of $\mathcal{K}$ are independent and spread over the KB. The minimum value is obtained for instance when all MISes of $\mathcal{K}$ share at least one formula. In this case, $I_{\mathcal{W}}(\mathcal{K}) = \sum_{1 \leq i \leq n} w_i$.

Alternatively, $I_{\mathcal{W}}$ can be expressed as the maximum value of $\sum_{p_i \in \mathcal{P}} (w_i \times I_{\mathcal{W}}(p_i))$, for all $\mathcal{P} \in \mathcal{P}_{MISes}(\mathcal{K})$. Note that $I_{\mathcal{W}}(p_i) = |p_i|$, since p_i is a closed set packing.

From the definition of $I_{\mathcal{W}}$, we can derive the following property expressing its relation with I_{MI} and $I_{\mathcal{CC}}$ measures.

Proposition 4 *Let $\mathcal{K}$ be a KB and $\mu = |MISes(\mathcal{K})| - I_{\mathcal{CC}}(\mathcal{K})$. Then, we have:*

$$(\sum_{i=2}^{\mu} w_i) + I_{\mathcal{CC}}(\mathcal{K}) \leq I_{\mathcal{W}}(\mathcal{K}) \leq I_{MI}(\mathcal{K}).$$

Proof: By definition of $I_{\mathcal{CC}}$, there exists a strong-partition $\langle D, R\rangle$ where $D = \{\mathcal{K}_1, \ldots, \mathcal{K}_{I_{cc}(\mathcal{K})}\}$ and $R = \mathcal{K} \setminus \cup_{\mathcal{K}' \in D}\mathcal{K}'$. Let us consider $\mathcal{P} = \{p_1, \ldots, p_\mu\}$ where $p_1 = \{M_1, \ldots M_{I_{cc}(\mathcal{K})}\}$ s.t. $M_i \in MISes(\mathcal{K}_i)$ for $1 \leq i \leq I_{\mathcal{CC}}(\mathcal{K})$, and $\{p_2, \ldots, p_\mu\} = MISes(\mathcal{K}) \setminus p_1$. Clearly, $\mathcal{P}$ is a c-partition of $\mathcal{K}$. According to Definition 12, it holds that $I_{\mathcal{W}}(\mathcal{K}) \geq I_{\mathcal{CC}}(\mathcal{K}) + \sum_{2 \leq i \leq \mu} w_i$. Also, we have $I_{\mathcal{W}}(\mathcal{K}) \leq I_{MI}(\mathcal{K})$ since $\{w_n\}_{n=1}^{+\infty}$ is a decreasing sequence and $w_1 = 1$. □

Interestingly, the weighted inconsistency measure $I_{\mathcal{W}}$ satisfies all the rational properties mentioned above.

Theorem 5 *$I_{\mathcal{W}}$ measure is a standard measure.*

Proof:

- `Consistency` and `Independence` are verified because $MISes(\mathcal{K}) \neq \emptyset$ iff $\mathcal{K}$ is inconsistent, and $MISes(\mathcal{K}) = MISes(\mathcal{K} \cup \{\alpha\})$ for $\alpha \in free(\mathcal{K} \cup \{\alpha\})$.
- `Monotonicity`: Let $\mathcal{K}$ and $\mathcal{K}'$ be two KBs, and $\mathcal{P} = \{p_1, \ldots, p_n\}$ a c-partition of $MISes(\mathcal{K})$. We have $MISes(\mathcal{K}) \subseteq MISes(\mathcal{K} \cup \mathcal{K}')$. Let $\{M'_1, \ldots, M'_m\} = MISes(\mathcal{K} \cup \mathcal{K}') \setminus MISes(\mathcal{K})$. Then, $\mathcal{P}' = \{p_1, \ldots, p_n, \{M'_1\}, \ldots, \{M'_m\}\}$ is a c-partition of $\mathcal{K} \cup \mathcal{K}'$. So, $\mathcal{W}(\mathcal{P}') \geq \mathcal{W}(\mathcal{P})$. Consequently, $I_{\mathcal{W}}(\mathcal{K} \cup \mathcal{K}') \geq I_{\mathcal{W}}(\mathcal{K})$.
- `MinInc`: For a single MIS M, there exists a unique c-partition $P = \{M\}$. Since $w_1 = 1$, $I_{\mathcal{W}}(M) = 1$.
- `Ind-decomposability`: The proof follows from the existence of a bijection between the partition of $\mathcal{K} \cup \mathcal{K}'$ and the union of the partitions of $\mathcal{K}$ and $\mathcal{K}'$.

□

Now we can compare different KBs by conflict degrees.

Definition 13 *Let $\mathcal{K}$ and $\mathcal{K}'$ be two KBs. We say that $\mathcal{K}$ is* less inconsistent *than $\mathcal{K}'$ iff $I_{\mathcal{W}}(\mathcal{K}) \leq I_{\mathcal{W}}(\mathcal{K}')$.*

Example 4 *Let us consider the following two KBs $\mathcal{K}_1 = \{a, \neg a, a \vee b, \neg b, b\}$, and $\mathcal{K}_2 = \{c, \neg c \wedge d, \neg d \wedge e \wedge f, \neg e, \neg f\}$. Given a sequence $w_i = \frac{1}{i}$, $i \in \mathbb{N}^*$,*

- *We have, $MISes(\mathcal{K}_1) = \{\{a, \neg a\}, \{b, \neg b\}, \{\neg a, a \vee b, \neg b\}\}$. Then, the maximal c-partition of $\mathcal{K}_1$ is $\mathcal{P} = \{p_1, p_2\}$ such that $p_1 = \{\{a, \neg a\}, \{b, \neg b\}\}$ and $p_2 = \{\neg a, a \vee b, \neg b\}$. It follows that $I_{\mathcal{W}}(\mathcal{K}_1) = |p_1| \times w_1 + |p_2| \times w_2 = 2 \times 1 + 1 \times \frac{1}{2} = \frac{5}{2}$.*
- *In a similar way, we obtain $I_{\mathcal{W}}(\mathcal{K}_2) = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{25}{12}$.*

As $I_{\mathcal{W}}(\mathcal{K}_2) < I_{\mathcal{W}}(\mathcal{K}_1)$, so $\mathcal{K}_2$ is less inconsistent than $\mathcal{K}_1$.

Next we study the relationship between the three standard measures $I_{\mathcal{W}}$, $I_{\mathcal{CC}}$, and I_{MI} for particular sequences of weights.

Proposition 6 *Let $\mathcal{K}$ be a KB and $\{w_n\}_{n=1}^{+\infty}$ a weighting sequence such that $w_1 = 1$ and $w_n = \lambda$ for all $n > 1$, where λ is a non negative constant satisfying $0 \leq \lambda \leq 1$. Then,*

$$I_{\mathcal{W}}(\mathcal{K}) = (1 - \lambda) \times I_{\mathcal{CC}}(\mathcal{K}) + \lambda \times I_{MI}(\mathcal{K}).$$

Proof: Note that $w_1 = 1$ and $w_n = \lambda$ for all $n > 1$, it is clear that a partition $\mathcal{P} = \{p_1, \ldots, p_n\}$ maximizing the $\mathcal{W}(\mathcal{P})$ must put in p_1 the maximal set of MISes that forms a closed set packing, which corresponds to $I_{\mathcal{CC}}$, and the remaining MISes can be distributed into the other elements p_i $(1 < i \leq n)$ with the same amount of contribution λ to $I_{\mathcal{W}}(\mathcal{K})$. So $I_{\mathcal{W}}(\mathcal{K}) = I_{\mathcal{CC}}(\mathcal{K}) + \lambda \times (I_{MI}(\mathcal{K}) - I_{\mathcal{CC}}(\mathcal{K})) = (1 - \lambda) \times I_{\mathcal{CC}}(\mathcal{K}) + \lambda \times I_{MI}(\mathcal{K})$. □

The following result is very important, because it reveals that our framework of weighted inconsistency measures is general enough to encompass some standard measures as specific cases, especially those given in (Hunter and Konieczny 2010; Jabbour, Ma, and Raddaoui 2014).

Theorem 7 *Let $\mathcal{K}$ be a KB. We have:*

$$I_{\mathcal{W}}(\mathcal{K}) = \begin{cases} I_{\mathcal{CC}}(\mathcal{K}) & \text{if } \lambda = 0 \\ \frac{I_{\mathcal{CC}}(\mathcal{K}) + I_{MI}(\mathcal{K})}{2} & \text{if } \lambda = \frac{1}{2} \\ I_{MI}(\mathcal{K}) & \text{if } \lambda = 1 \end{cases}$$

Proof: Direct consequence of Proposition 6. □

Corollary 8 *Let $\mathcal{K}$ be a KB and $\{w_n\}_{n=1}^{+\infty}$ a weighting sequence. Then, we have:*

$I_{\mathcal{W}}(\mathcal{K}) = I_{MI}(\mathcal{K})$ if and only if $\forall n \in [1, +\infty[, w_n = 1$.

Finally, let us stress that the definition of $I_{\mathcal{W}}$ is a general method to define an inconsistency measure. Following the same way, another measure can be obtained by defining the partitions by set packing instead of closed set packing (cf. Definition 6), which leads to the following new inconsistency value: $I_{\mathcal{SP}}(\mathcal{K}) = \max\{\mathcal{W}(\mathcal{P}) \mid \mathcal{P} \in \mathcal{P}_{MISes}(\mathcal{K})$ such that $\forall p \in \mathcal{P}$, p is a set packing$\}$.

Example 5 *Consider the following KBs $\mathcal{K}_1 = \{a, \neg a, a \vee b, \neg b, b\}$ and $\mathcal{K}_2 = \{a, \neg a \wedge b, \neg b \wedge c, \neg c\}$. Suppose that we use a sequence $w_i = \frac{1}{i}$, $i \in \mathbb{N}^*$. We have:*

- *$I_{\mathcal{W}}(\mathcal{K}_1) = I_{\mathcal{SP}}(\mathcal{K}_1) = 2 + \frac{1}{2} = \frac{5}{2}$,*
- *$I_{\mathcal{W}}(\mathcal{K}_2) = 1 + \frac{1}{2} + \frac{1}{3} = \frac{11}{6}$ and $I_{\mathcal{SP}}(\mathcal{K}_2) = 2 + \frac{1}{2} = \frac{5}{2}$.*

Note that both $\mathcal{K}_1$ and $\mathcal{K}_2$ have three MISes: $M^1_{\mathcal{K}_1} = \{a, \neg a\}$, $M^2_{\mathcal{K}_1} = \{\neg a, a \vee b, \neg b\}$, $M^3_{\mathcal{K}_1} = \{b, \neg b\}$, and $M^1_{\mathcal{K}_2} = \{a, \neg a \wedge b\}$, $M^2_{\mathcal{K}_2} = \{\neg a \wedge b, \neg b \wedge c\}$, $M^3_{\mathcal{K}_2} = \{\neg b \wedge c, \neg c\}$. Similarly, they both own two sized set packings, e.g. $\{M^1_{\mathcal{K}_1}, M^3_{\mathcal{K}_1}\}$ and $\{M^1_{\mathcal{K}_2}, M^3_{\mathcal{K}_2}\}$, respectively. However, their closed set packings are different, so are the c-partitions corresponding to $I_{\mathcal{W}}(\mathcal{K}_1)$ and $I_{\mathcal{W}}(\mathcal{K}_2)$: $\langle\{M^1_{\mathcal{K}_1}, M^3_{\mathcal{K}_1}\}, \{M^2_{\mathcal{K}_1}\}\rangle$ and $\langle\{M^1_{\mathcal{K}_2}\}, \{M^2_{\mathcal{K}_2}\}, \{M^3_{\mathcal{K}_2}\}\rangle$, respectively.

By their definitions, we have the following relation between $I_{\mathcal{SP}}$ and $I_{\mathcal{W}}$ measures:

Proposition 9 *Let $\mathcal{K}$ be a KB. The following inequality holds:*

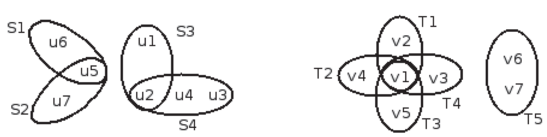
$$I_{\mathcal{SP}}(\mathcal{K}) \geq I_{\mathcal{W}}(\mathcal{K}).$$

Towards a Parameter-Free Standard Measure

An important property of the previously defined weight-based measure $I_{\mathcal{W}}$ lies in its ability to encompass some existing inconsistency measures, such as I_{MI} and $I_{\mathcal{CC}}$, as specific cases. However, such a measure is defined using an additional parameter w that must be tuned in practice. In the following, we will study the impact of the parameter w on the induced $I_{\mathcal{W}}$ value. Considering that an inconsistency measure is often used to rank different KBs, the goal is, therefore, to have a parameter-free measure that can return a meaningful ranking of different KBs based on their degrees of conflict.

First, we examine the ranking of different KBs under the $I_{\mathcal{W}}$ measure, illustrated by the following example.

Example 6 *Let us consider two KBs $\mathcal{K}_1 = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$ and $\mathcal{K}_2 = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$ whose MISes are depicted below by ellipses:*



We can see that some maximal c-partitions can be obtained (the left hand side for $\mathcal{K}_1$, the right hand side for $\mathcal{K}_2$):

So we have $V(\mathcal{K}_1) = \langle 2, 2\rangle$ and $V(\mathcal{K}_2) = \langle 2, 1, 1, 1\rangle$. Now, we take two different sequences $w_1^i = \frac{1}{i}$ and $w_2^i = \frac{1}{2^{i-1}}, i \in \mathbb{N}^*$. Then, for w_1^i we have $I_{\mathcal{W}}(\mathcal{K}_1) = 2\times 1 + 2\times \frac{1}{2} = 3$ and $I_{\mathcal{W}}(\mathcal{K}_2) = 2\times 1 + 1\times\frac{1}{2} + 1\times\frac{1}{3} + 1\times\frac{1}{4} = \frac{37}{12}$. Consequently, $I_{\mathcal{W}}(\mathcal{K}_1) < I_{\mathcal{W}}(\mathcal{K}_2)$. Meanwhile, if we consider the sequence w_2^i it holds that $I_{\mathcal{W}}(\mathcal{K}_1) = 2\times 1 + 2\times\frac{1}{2} = 3$ and $I_{\mathcal{W}}(\mathcal{K}_2) = 2\times 1 + 1\times\frac{1}{2} + 1\times\frac{1}{2^2} + 1\times\frac{1}{2^3} = \frac{23}{8}$, so $I_{\mathcal{W}}(\mathcal{K}_1) > I_{\mathcal{W}}(\mathcal{K}_2)$, which contradicts the previous ordering of $\mathcal{K}_1$ and $\mathcal{K}_2$ under w_1^i.

This example shows that there exist cases where different settings of the parameter w can lead to different rankings of KBs. To overcome this problem, we introduce below a new inconsistency measure, called I_{cf}, which is defined as a special *continued fraction* constructed from MIS partitions of a KB.

Definition 14 *Let $\mathcal{K}$ be a KB and $\mathcal{V}(\mathcal{K}) = \langle v_1, ..., v_n\rangle$ its conflict vector. The inconsistency measure I_{cf} is 0 if $\mathcal{V}(\mathcal{K})$ is a zero vector; otherwise I_{cf} is the continued fraction corresponding to the extended sequence $\mathcal{V}^e(\mathcal{K}) = \langle v_1, 1, v_2, 1, \ldots, v_{n-1}, 1, v_n\rangle$:*

$$I_{cf}(\mathcal{K}) = v_1 + \cfrac{1}{1 + \cfrac{1}{v_2 + \cfrac{1}{1 + \cfrac{1}{\ddots + \cfrac{1}{v_n}}}}}.$$

Note that the extra "1"s in the above definition are added into the vector to guarantee that Theorem 11 (below) holds.

Example 7 (Example 6 contd.) *From Definition 14, we have $I_{cf}(\mathcal{K}_1) = 2 + \frac{1}{1+\frac{1}{2}} = \frac{8}{3}$ and $I_{cf}(\mathcal{K}_2) = \frac{34}{13}$ whose corresponding extended sequence are $\mathcal{V}^e(\mathcal{K}_1) = \langle \mathbf{2}, 1, \mathbf{2}\rangle$ and $\mathcal{V}^e(\mathcal{K}_2) = \langle \mathbf{2}, 1, \mathbf{1}, 1, \mathbf{1}, 1, \mathbf{1}\rangle$. Then, we deduce that $\mathcal{K}_1$ is more inconsistent than $\mathcal{K}_2$, i.e., $\mathcal{K}_2$ is less conflicting than $\mathcal{K}_1$, under the measure I_{cf}.*

Similar to $I_{\mathcal{W}}$, we can see that $I_{cf}(\mathcal{K})$ satisfies all the desired properties of a standard measure.

Proposition 10 *I_{cf} measure is a standard measure.*

For any KBs $\mathcal{K}$ and $\mathcal{K}'$, suppose that their conflict vectors are $V(\mathcal{K}) = \langle v_1, \cdots, v_n\rangle$ and $V(\mathcal{K}') = \langle v_1', \cdots, v_m'\rangle$, respectively. Recall that we have a ranking between $\mathcal{K}$ and $\mathcal{K}'$ by their conflict vectors, that is, $\mathcal{K}$ is less conflicting than $\mathcal{K}'$ if $V(\mathcal{K}) \preceq V(\mathcal{K}')$.

Example 8 (Example 6 and 7 contd.) *Since $V(\mathcal{K}_1) = \langle 2, 2\rangle$ and $V(\mathcal{K}_2) = \langle 2, 1, 1, 1\rangle$, we have $V(\mathcal{K}_2) \preceq V(\mathcal{K}_1)$, that is, $\mathcal{K}_2$ is of less conflict than $\mathcal{K}_1$ according to their conflict vectors.*

The key property of the I_{cf} measure is characterized by the following proposition.

Theorem 11 *For any KBs $\mathcal{K}$ and $\mathcal{K}'$, $I_{cf}(\mathcal{K}) \leq I_{cf}(\mathcal{K}')$ if $\mathcal{K} \preceq \mathcal{K}'$.*

Proof: It is obvious that if $\mathcal{K} = \mathcal{K}'$, $I_{cf}(\mathcal{K}) = I_{cf}(\mathcal{K}')$. Now we suppose $\mathcal{K} \prec \mathcal{K}'$ and $V(\mathcal{K}) = \langle v_1, \cdots, v_n\rangle$, $V(\mathcal{K}') = \langle v_1', \cdots, v_m'\rangle$. Then, there exists $1 \leq k \leq min(n, m)$ such that $v_i = v_i'$ for $0 \leq i < k$ and $v_k < v_k'$. Therefore, all the subparts of the continued fraction of the form $\frac{1}{1+\ldots}$ are strictly smaller than 1. We deduce that $v_k + \frac{1}{1+\frac{1}{v_{k+1}+\ldots}} < v_k'$ by noting that v_k and v_k' are positive integers according to the definition of conflict vector. Consequently, $I_{cf}(\mathcal{K}) \leq I_{cf}(\mathcal{K}')$. □

Theorem 11 tells us that the ranking of KBs under the I_{cf} measure coincides with that defined by their conflict vectors.

On the Computation of Standard Measures

According to Corollary 3, computing the lower bound of standard measures is equivalent to computing a solution to a maximum closed set packing problem. In this section, we extend this result to the fine-grained measure $I_{\mathcal{W}}$ and I_{cf}. Indeed, we provide encodings for all the three measures into the Minimum Cost Satisfiability problem (MinCostSAT) (Miyazaki, Iwama, and Kambayashi 1996), which allows us to take a full advantage of the continuous progresses in practical SAT solving.

MinCostSAT-based Encoding of $I_{\mathcal{CC}}$

Let us recall that $I_{\mathcal{CC}}$ is the cardinality of the maximum closed set packing of $S = MISes(\mathcal{K})$ over the universe $U = \mathcal{K}$. Notice that a direct encoding of this problem can be obtained by modeling the conflicting sets of MISes, i.e., sets of MISes which are not closed set packings. Unfortunately, this approach is clearly inefficient because of the huge number of possible conflicting sets of MISes. To overcome this problem, we propose an original encoding that considers both S and U in MinCostSAT.

Definition 15 (MinCostSAT) *Let α be a CNF formula and f a cost function that associates a non-negative cost to each variable in $Var(\alpha)$. The* MinCostSAT *is the problem of finding a model $\mathcal{B}$ for α that minimizes the objective function:*

$$\mathcal{F}(\mathcal{B}) = \sum_{x \in Var(\alpha)} f(x) \times \mathcal{B}(x)$$

Let U be a universe and S a set of subsets of U. We associate a boolean variable X_e (resp. Y_{S_i}) to each $e \in U$ (resp. $S_i \in S$). The first formula allows us to only consider the pairwise disjoint subsets in S:

$$\bigwedge_{e \in U} \sum_{S_i \in S | e \in S_i} Y_{S_i} \leq 1 \tag{1}$$

The inequalities in (1) correspond to the AtMostOne constraint which is a special case of the well-known cardinality constraint. Several efficient encodings of the cardinality constraint to CNF have been proposed, most of which try to improve the efficiency of constraint propagation (e.g. (Bailleux and Boufkhad 2003; Sinz 2005)). In this case, the inequality $\sum_{S_i \in S | e \in S_i} Y_{S_i} \leq 1$ is encoded as follows using a sequential counter (Sinz 2005; Marques-Silva and Lynce 2007) (we suppose that $\sum_{S_i \in S | e \in S_i} Y_{S_i}$ can be rewritten as $\sum_{i=1}^{n} Y_{S_i}$):

$$\begin{gathered}(\neg Y_{S_1} \vee q_1) \wedge (\neg Y_{S_n} \vee \neg q_{n-1}) \\ \bigwedge_{1<i<n} ((\neg Y_{S_i} \vee q_i) \wedge (\neg q_{i-1} \vee q_i) \wedge (\neg Y_{S_i} \vee \neg q_{i-1}))\end{gathered} \tag{2}$$

where q_i is a fresh propositional variable for all $1 \leq i \leq n-1$.

The following formula is to express that $Y_{S_i} = 1$ if and only if for all $e \in S_i$, $X_e = 1$, i.e., $Y_{S_i} \Leftrightarrow (\sum_{e \in S_i} X_e = |S_i|)$:

$$\bigwedge_{S_i \in S} \bigwedge_{e \in S_i} \neg Y_{S_i} \vee X_e \tag{3}$$

$$\bigwedge_{S_i \in S} (Y_{S_i} \vee \bigvee_{e \in S_i} \neg X_e) \tag{4}$$

It is worth noticing that any solution to the inequalities (1), (3) and (4) represents a closed set packing of S.

In order to compute $I_{\mathcal{CC}}$, we need to maximize the sum $\sum_{S_i \in S} Y_{S_i}$. To encode this maximization problem as a MinCostSAT problem, we only need to rename each variable Y_{S_i} with $\neg Y'_{S_i}$ (Y'_{S_i} is a fresh propositional variable) in $\alpha = (1) \wedge (3) \wedge (4)$, for all $S_i \in S$. We note such renamed formula $\mathcal{R}(\alpha)$. The $MinCostSAT\text{-}I_{\mathcal{CC}}(U, S)$, encoding $I_{\mathcal{CC}}$, is defined as follows:

Problem: $MinCostSAT\text{-}I_{\mathcal{CC}}(U, S)$

$$\min \mathcal{F}(\mathcal{B}) = \sum_{x \in Var(\alpha)} f(x) \times \mathcal{B}(x)$$

subject to

$\mathcal{B} \in \mathcal{M}(\mathcal{R}(\alpha))$ with f defined as follows:

$\forall S_i \in S, f(Y'_{S_i}) = 1;$

$\forall X \in Var(\alpha) \setminus \{Y'_{S_i} \mid S_i \in S\}, f(X) = 0.$

MinCostSAT-based Encoding of $I_{\mathcal{W}}$

Similarly to the previous encoding of $I_{\mathcal{CC}}$, we provide a MinCostSAT-based encoding to compute $I_{\mathcal{W}}$ value. Indeed, we look for a partition of MISes such that each element of the partition is a closed set packing (see Definition 6).

Let $P = \{P_1, \ldots, P_n\}$ be a partition of S. Here, we associate a binary variable X_e^j to each element e in U to express that $e \in P_j$ and a set of binary variables $Y_{S_i}^j \in \{0, 1\}$ to each subset S_i in S to express that $S_i \in P_j$.

The first formula allows us to only consider the pairwise disjoint subsets in S w.r.t. P_j:

$$\bigwedge_{e \in U} \bigwedge_{j=1}^{n} \sum_{S_i \in S | e \in S_i} Y_{S_i}^j \leq 1 \tag{5}$$

The following formulas allow us to express that $Y_{S_i}^j = 1$ if and only if $\forall e \in S_i$, $X_e^j = 1$:

$$\bigwedge_{j=1}^{n} \bigwedge_{S_i \in S} \bigwedge_{e \in S_i} \neg Y_{S_i}^j \vee X_e^j \tag{6}$$

$$\bigwedge_{j=1}^{n} \bigwedge_{S_i \in S} (Y_{S_i}^j \vee \bigvee_{e \in S_i} \neg X_e^j) \tag{7}$$

The following formula expresses that each S_i has to be in exactly one element of $\mathcal{P}$:

$$\bigwedge_{S_i \in S} \sum_{j=1}^{n} Y_{S_i}^j = 1 \tag{8}$$

In order to obtain a sound encoding of $I_{\mathcal{W}}$, we have to maximize the following sum: $\sum_{j=1}^{n} (\sum_{S_i \in S} Y_{S_i}^j) \times w_j$.
A MinCostSAT encoding can be obtained by renaming the variables of the form $Y_{S_i}^j$ as $\neg Z_{S_i}^j$ ($Z_{S_i}^j$ is a fresh variable), in the same way as in our encoding of $I_{\mathcal{CC}}$. We use $\mathcal{R}'(\alpha)$ to denote the formula obtained from α by using this renaming where $\alpha = (5) \wedge (6) \wedge (7) \wedge (8)$.

Problem: $MinCostSAT\text{-}I_{\mathcal{W}}(U, S)$

$$\min \mathcal{F}(\mathcal{B}) = \sum_{x \in Var(\alpha)} f(x) \times \mathcal{B}(x)$$

subject to

$\mathcal{B} \in \mathcal{M}(\mathcal{R}'(\alpha))$ with f defined as follows:

$\forall S_i \in S$ and $\forall j \in \{1, \ldots, n\}, f(Z^j_{S_i}) = w_j$;

$\forall X \in Var(\alpha) \setminus \mathcal{Z}, f(X) = 0$

where $\mathcal{Z} = \{Z^j_{S_i} \mid S_i \in S, j \in \{1, \ldots, n\}\}$.

Computation of I_{cf}

Let us recall that the computation of I_{cf} is based on the conflict vector corresponding to a maximal c-partition w.r.t. a lexicographical ordering. The encoding of c-partitions is described by the constraints $\mathcal{C}^p$ =(5)∧ (6)∧ (7)∧ (8) used for computing $I_{\mathcal{W}}$. To calculate a maximal c-partition $\mathcal{P}$, we enumerate the models of $\mathcal{C}^p$ in a lexicographical ordering. Each time a model is found, a lex-constraint is added dynamically in order to search for a new model (c-partition) that is greater than the previous one w.r.t. $\preceq_{pr}$. In a second step, we derive a conflict vector associated to the maximal c-partition of $\mathcal{K}$, then we compute I_{cf} (see Definition 14).

Experiments

The previous encodings provide a way to benefit from the efficient SAT solvers for computing the family of standard inconsistency measures. In this section, we conduct a comparative evaluation of I_{MI}, $I_{\mathcal{CC}}$, $I_{\mathcal{W}}$, and I_{cf} measures. Our goal is to discover their strengths and complementarities. We implemented the algorithms based on an optimization-based SAT solver MiniSAT 2.2 (Eén and Sörensson 2003). The experiments were performed on a Xeon 3.2GHz (2 GB RAM) cluster with a timeout of one hour of CPU time based on the following three datasets:

Datasets 1 were taken from the MaxSAT competition[2] used in the context of MISes Enumeration (Previti and Marques-Silva 2013). These datasets encode real-world problems coming from rocket domain or automotive product configuration of car lines. For these instances, the state-of-the-art MISes enumerator *eMUS* (Previti and Marques-Silva 2013) was used to generate MISes of each KB. Note that *eMUS* is a partial MISes enumerator, so for certain instances, we only consider subsets of their MISes if *eMUS* cannot get all MISes before timeout.

Datasets 2 are instances from the error-tolerant reasoning (Ludwig and Peñaloza 2014) over large biomedical ontologies (SNOMED version 13.11d and the NCI thesaurus[3]) with their $\mathcal{EL}$++ versions (Baader, Brandt, and Lutz 2005). The data instances correspond to different brave and cautious consequences of the given ontologies and errors (Ludwig and Peñaloza 2014). The MISes for these data were computed by JUST tool (Ludwig 2014). We used only the instances treatable by JUST under its predefined timeout.

[2]http://www.satcompetition.org/2011/

[3]http://evs.nci.nih.gov/ftp1/NCI_Thesaurus

Datasets 3 contain randomly generated instances, named mpfs_m_n, having m MISes and each MIS is of the size n. The artificial datasets are designed because both *eMUS* and JUST can only terminate over instances whose MISes are highly correlated, which is reflected by the small $I_{\mathcal{CC}}$ values for the datasets 1 and 2 as shown in Table 1. The procedure is done by generating a random family of sets $\{S_1, \ldots, S_m\}$ to represent the set of MISes, with m varying between 50 and 200. Each S_i (corresponding to a MIS) is generated as a set of integers selected from the interval $[1, 100]$ to stand for formulas in S_i by their indexes.

Table 1 reports, for each instance, the inconsistency values under the four standard measures, i.e., I_{MI}, $I_{\mathcal{CC}}$, $I_{\mathcal{W}}$ (where $w_n = \frac{1}{n}$), and I_{cf}, from which we can draw the following conclusions:

The first interesting observation is that both $I_{\mathcal{W}}$ and I_{cf} metrics assign all instances with *distinct* inconsistency degrees. In contrast, many instances have equivalent $I_{\mathcal{CC}}$ or I_{MI} values. The $I_{\mathcal{CC}}$ is more problematic due to its small values of being either 1 or 2 in most cases for the first two real-world datasets. This means that the $I_{\mathcal{CC}}$ is not adequate in practice to be used as an improvement of I_{MI} to take into account inner structure of MISes while estimating inconsistencies. Meanwhile, it shows that our $I_{\mathcal{W}}$ and I_{cf} measures can be used to better distinguish different KBs according to their different inconsistency degrees, which benefits from the proposed MIS partition for a fine-grained analysis of the inner structure of MISes.

Secondly, compared to $I_{\mathcal{W}}$, the parameter-free measure I_{cf} has an important theoretical property that it guarantees the ordering defined over conflict vectors. However, from the experimental result, an advantage of the measure $I_{\mathcal{W}}$ in practice exists in a clear difference between the inconsistency values. For instance, $I_{\mathcal{W}}(Snomed\_typeI\_out2) = 6.53$ and $I_{\mathcal{W}}(Snomed\_typeI\_out48) = 4.16$ but their I_{cf} values are too close with a difference of less than 10^{-7}.

Finally, we note that there is a big difference between I_{MI} and the other measures (i.e. $I_{\mathcal{CC}}$, $I_{\mathcal{W}}$, I_{cf}). For example, the instances *C220_FV_RZ_13* and *apex_gr_2pin_w4.shuffled** from the datasets 1 have large I_{MI} values (6772 and 1500, respectively), but much smaller $I_{\mathcal{CC}}$, $I_{\mathcal{W}}$, and I_{cf} degrees ($\leq$ 10). Small values of $I_{\mathcal{CC}}$, $I_{\mathcal{W}}$, and I_{cf} indicate that the MISes of such instances are strongly interconnected, which can not be reflected by merely using the I_{MI} measure.

Related Work

In this section, we provide a brief overview of some works related to inconsistency measures.

To measure inconsistencies in classical logics, a range of logic-based approaches have been proposed. For instance, one can cite those based on probabilistic models (Knight 2002; Doder et al. 2010), multi-valued semantics (Grant 1978; Hunter 2002; 2003; Oller 2004; Hunter 2006; Grant and Hunter 2008; Ma et al. 2010; Xiao et al. 2010; Ma, Qi, and Hitzler 2011), maximal consistent subsets (Ammoura et al. 2015), minimal inconsistent subsets (Hunter and Konieczny 2008; Mu, Liu, and Jin 2011; 2012; Xiao and Ma 2012; Jabbour, Ma, and Raddaoui 2014) and hitting sets (Mu 2015;

Instance	#vars	#clauses	I_{MI}	$I_{\mathcal{CC}}$	$I_{\mathcal{W}}$	I_{cf}
C168_FW_UT_851	1909	6758	102	1	5.2	$\langle 1, 1, \cdots, 1\rangle_{102}$
C220_FV_RZ_13	1728	4014	6772	1	9.39	$\langle 1, 1, \cdots, 1\rangle_{6772}$
c880_gr_rcs_w5.shuffled	3280	44291	70	1	4.83	$\langle 1, 1, \cdots, 1\rangle_{70}$
rocket_ext.b	283	1844	75	1	4.9	$\langle 1, 1, \cdots, 1\rangle_{75}$
c7552-bug-gate-0*	2640	6989	1000	1	7.48	$\langle 1, 1, \cdots, 1\rangle_{1000}$
apex_gr_2pin_w4.shuffled*	1322	10940	1500	2	8.89	$\langle 1, 1, \cdots, 1\rangle_{1500}$
wb_conmax1.dimacs.filtered*	277950	1221020	20	2	4.54	2.73205080739
wb_4m8s4.dimacs.filtered*	463080	1759150	20	9	13.36	9.61803278689
wb1.dimacs.filtered*	49525	140091	71	50	59.16	50.61803398875
Snomed_typeI_out2	**SNOMED metrics**		112	2	6.53	2.73205080757
Snomed_typeI_out30	#axioms:	369194	8	1	3.78	$\langle 1, 1, \cdots, 1\rangle_{8}$
Snomed_typeII_out48	#concepts:	310013	19	2	4.16	2.73205080512
Snomed_typeII_out152	#roles:	58	21	2	4.54	2.73205080739
Snomed_typeII_out189			37	2	5.59	2.61803398875
NCI_typeI_out36	**NCI metrics**		41	1	4.96	$\langle 1, 1, \cdots, 1\rangle_{41}$
NCI_typeII_out157	#axioms:	159 805	79	2	5.2	2.61803398875
NCI_typeI_out233	#concepts:	104 087	82	1	5.96	$\langle 1, 1, \cdots, 1\rangle_{82}$
NCI_typeII_out425	#roles:	92	46	5	10.78	5.85410196625
mfsp_50_20	**mfsp_n_m**		50	5	11.12	5.82842712475
mfsp_100_50	n: # MISes		100	22	47.55	22.9582607431
mfsp_120_80	m: MIS size		120	20	46.03	20.9544511501
mfsp_150_60			150	11	29.50	11.9160797831
mfsp_200_50			200	11	34.79	11.9226162893

Table 1: Comparative Evaluation of I_{MI}, $I_{\mathcal{CC}}$, I_{cf}, and $I_{\mathcal{W}}$ with $w_n = \frac{1}{n}$. For I_{cf}, the results are either represented as the value or in the form of conflict vector $\langle 1, 1, \cdots, 1\rangle_n$ where n denotes the length of a conflict vector. Note that once $I_{\mathcal{CC}}(\mathcal{K}) = 1$, the conflict vector of $\mathcal{K}$ will be $\langle 1, 1, \cdots, 1\rangle_{|I_{MI}(\mathcal{K})|}$, which can be directly used for ranking KBs.

Thimm 2016).

Inconsistency measures based on a Shapley value are another alternative that exploits existing inconsistency measures to define a coalition-based game and then use the Shapley value to analyze the amount of inconsistency that can be imputed to each formula in a given KB (Hunter and Konieczny 2010). Recently, based on logical argumentation theory, another family of inconsistency measures for propositional logic has been proposed (Raddaoui 2015). Conspicuously, it is hardly possible to have a complete comparison of the existing measures. One way to categorize the existing metrics is with respect to their dependence on syntax or semantics. Semantic-based measures aim to compute the proportion of the language that is affected by the inconsistency. The inconsistency measures belonging to this class are often based on some paraconsistent semantics and, thus, syntax independent, because we can still find paraconsistent models for inconsistent KBs. Whilst, syntax-based approaches are concerned with the minimal number of formulas that cause inconsistencies. An overview of inconsistency measures for classical logics can be found in (Grant and Hunter 2011).

There is also related work on inconsistency measurement for quantitative logics. In particular, several works have extended existing inconsistency measures for classical frameworks to the probabilistic setting and investigates their properties. One can quote for example the family of inconsistency metrics, proposed by (Picado-Muiño 2011; Thimm 2013), based on the quantification of the minimal adjustments in the degrees of certainty (i.e., probabilities) of the statements necessary to make the KB consistent. In (Rodder and Xu 2001), another inconsistency measure for probabilistic conditional logic is proposed. It is based on generalized divergence which is a specific distance for probability functions.

Conclusion and Future Work

In this paper, based on the MIS partition, we have first presented an original framework for defining a family of inconsistency measures of KBs which encompasses several well-known measures as specific cases. Then, we simplified the framework to get a parameter-free measure which keeps a desired property in ranking inconsistent KBs. Moreover, the computational aspects of these new measures and of an existing lower bound of all standard measures have been explored using a MinCostSAT based encoding which enables the use of efficient SAT solvers. We have implemented the algorithms and tested them over real-world datasets. The preliminary but encouraging experimental results highlight that the new inconsistency measures can better distinguish different KBs by their conflict degrees in comparison to two well-known inconsistency metrics.

As future work, we plan to conduct further experimental validations of our proposed framework on application domains where inconsistency measure would be very helpful. Such domains include belief merging, argumentation and heterogeneous source integration and management.

Acknowledgments The second author benefits from the support of the ANR under the number ANR-15-CE23-0022.

References

Ammoura, M.; Raddaoui, B.; Salhi, Y.; and Oukacha, B. 2015. On measuring inconsistency using maximal consistent sets. In *ECSQARU*, 267–276.

Baader, F.; Brandt, S.; and Lutz, C. 2005. Pushing the $\mathcal{EL}$ envelope. In *IJCAI*, 364–369.

Bail, S.; Horridge, M.; Parsia, B.; and Sattler, U. 2011. The justificatory structure of the NCBO bioportal ontologies. In *ISWC*, 67–82.

Bailleux, O., and Boufkhad, Y. 2003. Efficient CNF encoding of boolean cardinality constraints. In *CP*, 108–122.

Barragans-Martinez, A. B.; Arias, J. P.; and Vilas, A. F. 2004. On measuring levels of inconsistency in multi-perspective requirements specifications. In *PRISE*, 21–30.

Benferhat, S.; Dubois, D.; and Prade, H. 1995. A local approach to reasoning under inconsistency in stratified knowledge bases. In *ECSQARU*, 36–43.

Besnard, P. 2014. Revisiting postulates for inconsistency measures. In *JELIA*, 383–396.

Doder, D.; Raskovic, M.; Markovic, Z.; and Ognjanovic, Z. 2010. Measures of inconsistency and defaults. *Int. J. Approx. Reasoning* 51(7):832–845.

Eén, N., and Sörensson, N. 2003. An extensible SAT-solver. In *SAT*, 502–518.

Grant, J., and Hunter, A. 2006. Measuring inconsistency in knowledgebases. *J. Intell. Inf. Syst.* 27(2):159–184.

Grant, J., and Hunter, A. 2008. Analysing inconsistent first-order knowledgebases. *Artif. Intell.* 172(8-9):1064–1093.

Grant, J., and Hunter, A. 2011. Measuring the good and the bad in inconsistent information. In *IJCAI*, 2632–2637.

Grant, J., and Hunter, A. 2013. Distance-based measures of inconsistency. In *ECSQARU*, 230–241.

Grant, J. 1978. Classifications for inconsistent theories. *Notre Dame Journal of Formal Logic* 19(3):435–444.

Hunter, A., and Konieczny, S. 2008. Measuring inconsistency through minimal inconsistent sets. In *KR*, 358–366.

Hunter, A., and Konieczny, S. 2010. On the measure of conflicts: Shapley inconsistency values. *Artif. Intell.* 174(14):1007–1026.

Hunter, A.; Parsons, S.; and Wooldridge, M. 2014. Measuring inconsistency in multi-agent systems. *Unstliche Intelligenz* 28(3):169–178.

Hunter, A. 2002. Measuring inconsistency in knowledge via quasi-classical models. In *AAAI*, 68–73.

Hunter, A. 2003. Evaluating significance of inconsistencies. In *IJCAI*, 468–478.

Hunter, A. 2006. How to act on inconsistent news: Ignore, resolve, or reject. *Data Knowl. Eng.* 57(3):221–239.

Jabbour, S.; Ma, Y.; Raddaoui, B.; and Saïs, L. 2014. Prime implicates based inconsistency characterization. In *ECAI*, 1037–1038.

Jabbour, S.; Ma, Y.; Raddaoui, B.; Sais, L.; and Salhi, Y. 2015. On structure-based inconsistency measures and their computations via closed set packing. In *AAMAS*, 1749–1750.

Jabbour, S.; Ma, Y.; and Raddaoui, B. 2014. Inconsistency measurement thanks to MUS-decomposition. In *AAMAS*, 877–884.

Kalyanpur, A.; Parsia, B.; Sirin, E.; and Hendler, J. A. 2005. Debugging unsatisfiable classes in OWL ontologies. *J. Web Sem.* 3(4):268–293.

Knight, K. 2002. Measuring inconsistency. *J. Philosophical Logic* 31(1):77–98.

Ludwig, M., and Peñaloza, R. 2014. Brave and cautious reasoning in EL. In *DL*, 274–286.

Ludwig, M. 2014. Just: a tool for computing justications w.r.t. EL ontologies. In *ORE*.

Ma, Y.; Qi, G.; Xiao, G.; Hitzler, P.; and Lin, Z. 2010. Computational complexity and anytime algorithm for inconsistency measurement. *Int. J. Software and Informatics* 4(1):3–21.

Ma, Y.; Qi, G.; and Hitzler, P. 2011. Computing inconsistency measure based on paraconsistent semantics. *J. Log. Comput.* 21(6):1257–1281.

Marques-Silva, J. P., and Lynce, I. 2007. Towards robust cnf encodings of cardinality constraints. In *CP*, 483–497.

Miyazaki, S.; Iwama, K.; and Kambayashi, Y. 1996. Database queries as combinatorial optimization problems. In *CODAS*, 477–483.

Mu, K.; Liu, W.; Jin, Z.; and Bell, D. A. 2011. A syntax-based approach to measuring the degree of inconsistency for belief bases. *Int. J. Approx. Reasoning* 52(7):978–999.

Mu, K.; Liu, W.; and Jin, Z. 2011. A general framework for measuring inconsistency through minimal inconsistent sets. *Knowl. Inf. Syst.* 27(1):85–114.

Mu, K.; Liu, W.; and Jin, Z. 2012. Measuring the blame of each formula for inconsistent prioritized knowledge bases. *J. Log. Comput.* 22(3):481–516.

Mu, K. 2015. Responsibility for inconsistency. *International Journal of Approximate Reasoning* 61:43–60.

Oller, C. A. 2004. Measuring coherence using lp-models. *J. Applied Logic* 2(4):451–455.

Picado-Muiño, D. 2011. Measuring and repairing inconsistency in probabilistic knowledge bases. *Int. J. Approx. Reasoning* 52(6):828–840.

Previti, A., and Marques-Silva, J. 2013. Partial MUS enumeration. In *AAAI*.

Qi, G.; Liu, W.; and Bell, D. A. 2005. Measuring conflict and agreement between two prioritized belief bases. In *IJCAI*, 552–557.

Raddaoui, B. 2015. Computing inconsistency using deductive argumentation. In *ICAART*, 164–172.

Rodder, W., and Xu, L. 2001. Elimination of inconsistent knowledge in the probabilistic expert system-shell SPIRIT (in German). In *Operations Research - Springer-Verlag*, 260–265.

Sinz, C. 2005. Towards an optimal cnf encoding of boolean cardinality constraints. In *CP*, 827–831.

Thimm, M. 2013. Inconsistency measures for probabilistic logics. *Artif. Intell.* 197:1–24.

Thimm, M. 2016. Stream-based inconsistency measurement. *Int. J. Approx. Reasoning* 68:68–87.

Xiao, G., and Ma, Y. 2012. Inconsistency measurement based on variables in minimal unsatisfiable subsets. In *ECAI*, 864–869.

Xiao, G.; Lin, Z.; Ma, Y.; and Qi, G. 2010. Computing inconsistency measurements under multi-valued semantics by partial Max-SAT solvers. In *KR*.