International Journal of Approximate Reasoning ••• (••••) •••–•••

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Quantifying conflicts in propositional logic through prime implicates

Said Jabbour [a], Yue Ma [b,*], Badran Raddaoui [c,*], Lakhdar Sais [a]

[a] CRIL – CNRS UMR 8188, University of Artois, F-62307 Lens Cedex, France
[b] LRI, Univ. Paris-Sud, CNRS, Université Paris-Saclay, F-91405 Orsay Cedex, France
[c] SAMOVAR, Télécom SudParis, CNRS, Univ. Paris-Saclay, F-91011 Evry Cedex, France

A R T I C L E   I N F O

Article history:
Received 13 October 2015
Received in revised form 28 December 2016
Accepted 30 December 2016
Available online xxxx

Keywords:
Knowledge representation
Inconsistency measure
Propositional logic
Conflicting variables
Prime implicates

A B S T R A C T

Quantifying conflicts is recognized as an important issue for handling inconsistencies. Indeed, an inconsistency measure can be employed to support knowledge engineers in building a consistent and usable knowledge base or providing insights on how to repair an inconsistent one. Good measures are supposed to satisfy a set of rational properties. However, defining sound properties is sometimes problematic. In this paper, we emphasize one such property, named *dominance*, rarely satisfied by syntactic measures. Based on prime implicates canonical representation, we first introduce the notion of conflicting variable and use it to refine an existing inconsistency measure defined by minimally unsatisfiable sets (MUSes). Then, we provide a semantics characterization allowing us to establish relationships with multi-valued semantics. Secondly, we propose a new measure based on the notion of deduced MUSes (DMUSes), to circumscribe the internal conflicts in a given knowledge base. We also prove that this measure satisfies a new but weaker form of dominance. Finally, we show how inconsistency measures based on hitting sets of minimal inconsistent sets can be extended using hitting sets of DMUSes.

1. Introduction

Inconsistencies are often encountered, especially when information come from different sources. Reasoning on inconsistent knowledge bases is problematic if one wants to use classical inference rules. In this view, consistency should be recovered by removing incorrect pieces of information. This may be done, for instance by analyzing and quantifying the amount of conflict of the set of contradictory information. Measuring conflicts has gained considerable attention in the field of Artificial Intelligence [1]. It is of particular importance for comparing different knowledge bases by their inconsistency levels [2]. It was also proved useful and attractive in diverse scenarios, including software specifications [3], e-commerce protocols [4], belief merging [5], news reports [6], integrity constraints [7], requirements engineering [3], databases [8,9], semantic web [10], and network intrusion detection [11].

Several logic-based inconsistency measures have been studied and there are different ways to categorize them. One way is by their dependence on the language or formula: the former aims to compute the proportion of the language affected by inconsistency [2,12,13,6,14–19]. Whilst, the latter is concerned with the minimal number of formulas that cause inconsistencies, often through minimal unsatisfiable subsets [20–22,9,23]. Some other measures are based on both [24,25]. These different measures can also be classified by being formula or knowledge base oriented. For example, the inconsistency mea-

* Corresponding authors.
E-mail addresses: jabbour@cril.fr (S. Jabbour), ma@lri.fr (Y. Ma), badran.raddaoui@telecom-sudparis.eu (B. Raddaoui), sais@cril.fr (L. Sais).

http://dx.doi.org/10.1016/j.ijar.2016.12.017

sures proposed in [24,25] consist in quantifying the contribution of a formula to the inconsistency of the whole knowledge base containing it, while the other mentioned measures aim to quantify the inconsistency of the whole knowledge base. Furthermore, some established basic properties [25] such as *consistency, monotony, and free formula independence* are proposed to evaluate the quality of inconsistency measures.

In this paper, we focus on knowledge base oriented inconsistency measures, from both the language and the formula aspects. Moreover, we investigate a particular property, namely *dominance*, which is often problematic for formula-based inconsistency measures. The aim is to investigate novel language- and formula-based inconsistency measures while answering the limitations of existing ones by satisfying the desired properties.

Inspired by the example given in [18], suppose that there are n groups of people polled on a set of policies $\{p_1, ..., p_m\}$. The poll result of each group is formalized in propositional logic. For example, $\{p_1 \wedge \neg p_2, p_1 \vee p_3\}$ expresses that in this group there's one voter who votes for p_1 but against p_2, and the other voter supports either p_1 or p_3. Now consider the poll results of two groups $\gamma_1 = \{p_1 \wedge p_2, \neg p_2\}$ and $\gamma_2 = \{p_1, \neg p_1 \vee p_2, \neg p_2, p_2\}$, which are both classically inconsistent. Then, we can use different measures to compare γ_1 and γ_2 in terms of their inconsistency degrees. For instance, the measure ID_4 [6] ensures that γ_1 contains one unit of inconsistency, which seems reasonable because the conflict is merely on p_2 within this group, but this measure treats γ_2 equivalently even though there are indeed two conflicting subgroups: $\{p_1, \neg p_1 \vee p_2, \neg p_2\}$ and $\{\neg p_2, p_2\}$. In contrast, the measure ID_{MUS} [18] considers that both poll results have two units of inconsistency because there are two literals (p_1 and p_2) involved in at least one subgroup with conflicts. In short, the measure ID_4 ignores certain inconsistencies in γ_2 and ID_{MUS} overestimates inconsistency in γ_1, so does ID_Q [12] since ID_Q is always equal or larger than ID_{MUS}.

To improve these language-based measures, we propose a new notion, called *conflicting variable*, from which we derive an inconsistency measure ID^c_{MUS} that can properly distinguish the above voting results γ_1 and γ_2. Compared with ID_4 and ID_Q, the MUS based measure I_{MI} can distinguish γ_1 and γ_2. However, as argued in [26], I_{MI} does not satisfy the dominance property. In this paper, we further refine the notion of MUS and propose a new one called *deduced MUS* (*DMUS* for short) which leads to an interesting inconsistency measure that satisfies a new restrictive but more intuitive dominance property, called *weak dominance*. Our proposed framework makes use of prime implicates, a canonical representation of boolean formulas. This representation allows us to consider more finely the contributions of each formula to inconsistency, and to avoid redundancies (or equivalent conflicts). Notice that a similar requirement has also been identified in the context of ontological entailment justifications [27], where the proposed approach breaks down the axiom into atomic structures in order to identify the causes of unsatisfiability.

The paper is organized as follows: Section 2 gives necessary preliminaries and describes popular inconsistency measures relevant to the present work. In Section 3, we introduce a new notion, named *conflicting variable*, and discuss the computational complexity for related problems. In Section 4, we apply this notion to define a new inconsistency measure ID^c_{MUS} and prove that it is a refinement located between ID_4 and ID_{MUS}, that is, $ID_4 \leq ID^c_{MUS} \leq ID_{MUS}$. To elaborate formula-based inconsistency measures, we propose in Section 5 the MUS based logical deduction (DMUS) and use it to define a new measure, denoted by I_{DM}. We also show that I_{DM} is more interesting than I_{MI} because it can fulfill the *weak dominance* property. In Section 6, we show how inconsistency measures based on hitting sets of the minimal inconsistent sets can be extended using hitting sets of DMUSes. We provide some review of related work in Section 7 before concluding with some perspectives.

2. Preliminaries

Throughout this paper, we consider the propositional language $\mathcal{L}$ built over a countable infinite set of propositional symbols $\mathcal{P}$ using classical logical connectives $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow\}$. We will use letters such as p and q to denote propositional variables, and Greek letters like α and β to denote propositional formulas. The symbols $\top$ and $\bot$ denote *tautology* and *contradiction*, respectively. A literal is either a propositional variable (p) or its negation ($\neg p$), and a clause is a disjunction (possibly written as a set) of literals. A CNF formula is a conjunction of clauses $c_1 \wedge c_2 \wedge \cdots \wedge c_n$, also represented as a set of clauses $\{c_1, c_2, \ldots, c_n\}$ for simplicity. In contrast, a DNF is a disjunction of conjunctions of literals. For a set S, $|S|$ denotes its cardinality.

A *knowledge base* K consists of a finite set of propositional formulas. K is a CNF knowledge base, if it is a set of clauses. We denote by $Var(K)$ the set of variables occurring in K. Further, K is inconsistent if there is a formula α in the language $\mathcal{L}$ such that $K \vdash \alpha$ and $K \vdash \neg\alpha$, where $\vdash$ is the deduction in classical propositional logic. In this paper, we assume that K contains only consistent formulas, following Grant and Hunter [9]. If K is inconsistent, a *Minimal Unsatisfiable Subset (MUS)* of K is defined as follows:

Definition 1 *(MUS).* Let K be a knowledge base and $M \subseteq K$. M is a Minimal Unsatisfiable Subset (MUS) of K if $M \vdash \bot$ and $\forall M' \subsetneq M$, $M' \nvdash \bot$.

We write *MUSes*(K) to denote the set of minimal inconsistent subsets of K. Obviously, an inconsistent knowledge base K can have multiple minimal inconsistent subsets. If *MUSes*$(K) = \{K\}$, we call K itself a MUS. In particular, a CNF formula α is a MUS if $\alpha \vdash \bot$ but $\alpha \setminus \{c\} \nvdash \bot$ for any clause $c \in \alpha$. A formula α that is not involved in any minimal inconsistent set of K is called *free formula*. The set of free formulas of K is written $free(K) = \{\alpha \mid \nexists\, M \in MUSes(K) \text{ s.t. } \alpha \in M\}$.