International Journal of Approximate Reasoning 61 (2015) 43–60

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Responsibility for inconsistency

Kedian Mu

School of Mathematical Sciences, Peking University, Beijing 100871, PR China

ARTICLE INFO

Article history:
Received 4 August 2014
Received in revised form 20 April 2015
Accepted 24 April 2015
Available online 28 April 2015

Keywords:
Inconsistency
Responsibility
Causality models
Minimal inconsistent subsets
Minimal correction subsets

ABSTRACT

It is desirable to identify the degree of responsibility of each part of a knowledge base for the inconsistency of that base to make some necessary trade-off decisions on restoring the consistency of that base. In this paper, we propose a measurement for the degree of responsibility of each formula in a knowledge base for the inconsistency of that base. This measurement is given in terms of minimal inconsistent subsets of a knowledge base. Moreover, it can be well explained in the context of causality and responsibility presented by Chockler and Halpern [1].

1. Introduction

It seems inevitable that inconsistencies will occur when information is gathered from heterogeneous or distributed sources, especially at the advent of the Web. Techniques for handling inconsistency have been given much attention in the artificial intelligence community. In particular, it has been increasingly recognized that measurements of inconsistency for knowledge bases play an important role in analyzing inconsistent knowledge bases in many applications such as requirements engineering [2,3], expert systems in medicine [4], belief change [5], merging knowledge bases [6], ontology management [7] and intrusion detections in security [8].

A number of inconsistency measures have been proposed so far, including the maximal η-consistency [9,10] and the strict n-consistency [11], measures based on variables or paraconsistent models [12–17], measures based on the cost of test actions [18], measures based on minimal inconsistent subsets [19,20,16,21–23], measures based on minimal inconsistent subsets and consistent subsets [24], measures based on minimal proofs [25], measures based on prime implicates [26], the Shapley inconsistency value [5], measures for stratified knowledge bases [27], and measures for probabilistic knowledge bases [4,28–30].

Most of these inconsistency measures focus on assessing the degree of inconsistency for a whole knowledge base. However, in many practical applications, we are more interested in assessing the degree of responsibility of each formula in a knowledge base for the inconsistency of that base in the case that we want to resolve the inconsistency by changing some formulas involved in it.

Intuitively, identifying the degree of responsibility of a given formula in a knowledge base for the inconsistency of that base should not only depend on how to characterize the inconsistency, but also on how to formulate the causation of the inconsistency. A desirable measurement for the degree of responsibility of each formula in a knowledge base for the inconsistency of that base should be well explained from a causal perspective. That is, a formula has to bear some responsibility for the inconsistency only if the formula can be definitely shown as a cause of the inconsistency in some

E-mail address: mukedian@math.pku.edu.cn.

http://dx.doi.org/10.1016/j.ijar.2015.04.007
0888-613X/

formulation of causality of inconsistency. Moreover, its degree of responsibility should grasp some intended meaning of the formula being a cause of the inconsistency.

However, causality is a subtle topic in philosophy. There have been many attempts to define causality in literature from Hume to the present [1]. The *counterfactual dependence* is considered as a common ground of these attempts. Roughly speaking, A counterfactually depends on B if it is the case that if B had not happened, then A would not have happened. Recently, Halpern and Pearl's structural causal model has been influential in the computer science research community because this model extracts the generally accepted aspects of causality into a rigorous definition [31]. Roughly speaking, Halpern and Pearl's structural causal model presented in [32] considers that A is a cause of B if A counterfactually depends on B under some *contingency*. Essentially, the introduction of contingency makes the structural casual model more useful to capture the subtlety of causality than many other models based on counterfactual dependence. Moreover, it plays an important role in Chockler and Halpern's notion of responsibility presented in [1], which is an extension of Halpern and Pearl's definition of causality. To be more precise, Chockler and Halpern defined the degree of responsibility of A for B as $\frac{1}{N+1}$ if A is a cause of B, where N is the minimal number of changes that have to be made to obtain a contingency where B counterfactually depends on A (If A is not a cause of B, then the degree of responsibility is 0.) [1]. Then from the point of view of structural causal models, a formula of a knowledge base would be considered as a cause of the inconsistency in that base only if we can find a contingency where the inconsistency counterfactually depends on the formula. Moreover, its degree of responsibility for the inconsistency should depend on the minimal number of changes that have to be made to obtain such a contingency.

Inspired by these notions, in this paper, we propose a measurement for the degree of responsibility of each formula in a knowledge base for the inconsistency of that base ("the degree of responsibility of each formula for inconsistency" for short). Informally speaking, we use minimal inconsistent subsets of a knowledge base to characterize the inconsistency in that base. Here a minimal inconsistent subset of a knowledge base refers to an inconsistent subset of that base such that none of its proper subsets is inconsistent. Then the degree of responsibility of a formula for inconsistency is defined as 0 if the formula does not belong to any minimal inconsistent subset of that base. Otherwise, we define the degree of responsibility of that formula for inconsistency as $\frac{1}{1+N}$, where N is the minimal number of formulas that have to be removed from K in order to break all the minimal inconsistent subsets not containing the formula under the constraint that at least one minimal inconsistent subset containing the formula remains unbroken. Furthermore, we show that this measurement could be well explained in the context of causality and responsibility presented by Chockler and Halpern [1]. In detail, we show that a formula is a cause of the inconsistency if and only if it belongs to at least one minimal inconsistent subset when we formulate the causation of inconsistency by using Halpern and Pearl's causal model, moreover, the measurement built upon minimal inconsistent subsets does indeed capture the degree of responsibility in the causal model. In addition, allowing for the duality of minimal inconsistent subsets and minimal correction subsets [33] (here a minimal correction subset of a knowledge base refers to a subset-minimal set of formulas which need to be removed from the base in order to restore the consistency of the base), we give the measurement in terms of minimal correction subsets alternatively.

The rest of this paper is organized as follows. In Section 2 we provide some necessary notions about inconsistency as well as knowledge bases. Then we introduce Halpern and Pearl's definition of causality presented in [32] and Chockler and Halpern's notion of responsibility presented in [1], respectively. In Section 3 we propose a measurement of the degree of responsibility of each formula of a knowledge base for the inconsistency of that base. In Section 4, we give an explanation of this measurement by formulating the problem of inconsistency in the context of causal model presented by Chockler and Halpern [1]. In Section 5 we study logical properties and computational complexity issues of this measurement. In Section 6 we compare our approach with some closely related work. In Section 7, we present an example to illustrate the application of the measurement in the domain of requirements engineering. Finally, we conclude this paper in Section 8.

2. Preliminaries

2.1. Knowledge base and inconsistency

Throughout this paper, we use a finite propositional language. Let $\mathcal{P}$ be a finite set of propositional symbols (atoms) and $\mathcal{L}$ a propositional language built from $\mathcal{P}$ under connectives $\{\neg, \wedge, \vee\}$. We use a, b, c, $\cdots$ to denote the propositional symbols (variables), and $\alpha, \beta, \gamma, \cdots$ to denote the propositional formulas.

A knowledge base K is a finite set of propositional formulas. Just for simplicity, we assume that each knowledge base is non-empty. K is inconsistent if there is a formula α such that $K \vdash \alpha$ and $K \vdash \neg\alpha$, where $\vdash$ is the classical consequence relation. We abbreviate $\alpha \wedge \neg\alpha$ as $\perp$ when there is no confusion. Then an inconsistent knowledge base K is denoted by $K \vdash \perp$. In particular, we abbreviate $\{\alpha\} \vdash \perp$ as $\alpha \vdash \perp$, and say α is *inconsistent* or *self-contradictory*.

Moreover, an inconsistent knowledge base K is called a *minimal inconsistent set* if none of its proper subsets is inconsistent. If $S \subseteq K$ and S is a minimal inconsistent set, then we call S a *minimal inconsistent subset*(or *minimal unsatisfiable subset*) of K [20]. We use $\mathsf{MI}(K)$ to denote the set of all the minimal inconsistent subsets of K, i.e.,

$$\mathsf{MI}(K) = \{S \subseteq K \mid S \vdash \perp \text{ and } \forall S' \subset S, S' \nvdash \perp\}.$$

In addition, we abbreviate $\bigcup_{S \in \mathsf{MI}(K)} S$ and $\bigcap_{S \in \mathsf{MI}(K)} S$ as $\bigcup \mathsf{MI}(K)$ and $\bigcap \mathsf{MI}(K)$, respectively.

In syntax-based applications, $\mathsf{MI}(K)$ could be used to characterize the inconsistency of K, since one needs to remove only one formula from each minimal inconsistent subset to resolve the inconsistency [34,20].

A formula in K is called a *free formula* if this formula does not belong to any minimal inconsistent subset of K [20]. That is, free formulas have nothing to do with minimal inconsistent subsets of K. We use $\mathsf{FREE}(K)$ to denote the set of all the free formulas of K, i.e.,

$$\mathsf{FREE}(K) = \{\alpha \in K \mid \text{for all } S \in \mathsf{MI}(K), \alpha \notin S\}.$$

On the other hand, *minimal correction subsets* are also used to characterize the inconsistency of a knowledge base [35]. A subset R of K is called a *minimal correction subset* of K if $K \setminus R \nvdash \bot$ and for any proper subset R' of R, $K \setminus R' \vdash \bot$. We use $\mathsf{MC}(K)$ to denote the set of all the minimal correction subsets of K.

2.2. Causality and responsibility

Here we give introductions to Halpern and Pearl's causal model [32] and Chockler and Halpern's notion of responsibility [1], respectively. The material is largely taken from Section 2 and Section 3 in [1].

The variables involved in Halpern and Pearl's causal model are classified into two classes, namely exogenous variables and endogenous variables [32]. As explained in [1], values of each exogenous variable involved in a given model are determined by factors outside the model, while values of each endogenous variable involved in the model are ultimately determined by the exogenous variables.

A signature is a tuple $\mathcal{S} = \langle \mathcal{U}, \mathcal{V}, \mathcal{R} \rangle$, where $\mathcal{U}$ is a finite set of exogenous variables, $\mathcal{V}$ is a finite set of endogenous variables, and $\mathcal{R}$ associates with every variable $Y \in \mathcal{U} \cup \mathcal{V}$ a finite nonempty set $\mathcal{R}(Y)$ of possible values for Y [1,32].

A causal model over signature $\mathcal{S}$ is a tuple $M = \langle \mathcal{S}, \mathcal{F} \rangle$, where $\mathcal{F}$ associates with every endogenous variable $X \in \mathcal{V}$ a function F_X such that $F_X : ((\times_{U \in \mathcal{U}} \mathcal{R}(U)) \times (\times_{Y \in \mathcal{V} \setminus \{X\}} \mathcal{R}(Y))) \to \mathcal{R}(X)$ [1,32]. In particular, M is called a *binary causal model* if $\mathcal{R}(Y)$ contains only two values for each $Y \in \mathcal{U} \cup \mathcal{V}$ [1,32].

As explained in [1,32], F_X describes how the value of the endogenous variable X is determined by the values of all other variables in $\mathcal{U} \cup \mathcal{V}$. Then processes of assigning values to variables in M are characterized by the equations determined by $\{F_X : X \in \mathcal{V}\}$. On the other hand, if we set a value for each exogenous variable, these equations also provide counterfactual information. Here we use the following example taken from [1] to illustrate these explanations. Suppose that $F_X(Y, Z, U) = U + Y$ ($X = U + Y$ for short) and the value of the exogenous variable $U = 2$, then if $Y = 3$, then $X = 5$. On the other hand, if the value of Y were forced to be 4, then the value of X would be 6, regardless of what value X and Z actually take in the real world.

Given a causal model M, its causal network is a directed graph with nodes corresponding to the random variables in $\mathcal{V}$ and an edge from a node labeled X to one labeled Y if F_Y depends on the value of X [1,32]. If the associated causal network of M is a directed acyclic graph, then we call M *a recursive model* [1,32]. It has been stated in [1] that if M is a recursive causal model, then there is always a unique solution to the equations in M, given a setting for the variables in $\mathcal{U}$. We are more interested in *binary recursive causal models* in this paper.

We use $\vec{X}$ and $\vec{x}$ to denote a (possibly empty) vector of variables in $\mathcal{V}$ and values for the variables in $\vec{X}$, respectively. We use $\vec{X} \leftarrow \vec{x}$ to denote the case of setting the values of the variables in $\vec{X}$ to $\vec{x}$. We use $\vec{u}$ to denote a setting for the variables in $\mathcal{U}$. Here we call $\vec{u}$ a context [1,32].

Given $\vec{X} \leftarrow \vec{x}$, a new causal model denoted $M_{\vec{X} \leftarrow \vec{x}}$ over the signature $\mathcal{S}_{\vec{X}} = \langle \mathcal{U}, \mathcal{V} - \vec{X}, \mathcal{R}|_{\mathcal{V} - \vec{X}} \rangle$, is defined as $M_{\vec{X} \leftarrow \vec{x}} = \langle \mathcal{S}_{\vec{X}}, \mathcal{F}^{\vec{X} \leftarrow \vec{x}} \rangle$, where $F_Y^{\vec{X} \leftarrow \vec{x}}$ is obtained from F_Y by setting the values of the variables in $\vec{X}$ to $\vec{x}$ [1,32]. For example, suppose that $F_Y(X, Z, U) = X + Z + U$, then $F_Y^{X \leftarrow 3} = 3 + Z + U$.

Given a signature $\mathcal{S} = \langle \mathcal{U}, \mathcal{V}, \mathcal{R} \rangle$, a primitive event is a formula of the form $X = x$, where $X \in \mathcal{V}$ and $x \in \mathcal{R}(X)$ [1,32]. In general, for $\vec{X} = (X_1, X_2, \cdots, X_n)$ and $\vec{x} = (x_1, x_2, \cdots, x_n)$, we abbreviate $(X_1 = x_1) \wedge (X_2 = x_2) \wedge \cdots \wedge (X_n = x_n)$ as $\vec{X} = \vec{x}$.

A basic causal formula defined in [1,32] is in the form of

$$[Y_1 \leftarrow y_1, \cdots, Y_k \leftarrow y_k]\varphi,$$

where

- φ is a Boolean combination of primitive events;
- $Y_1, \cdots, Y_k$ are distinct variables in $\mathcal{V}$; and
- $y_i \in \mathcal{R}(Y_i)$.

As explained in [1,32], $[Y_1 \leftarrow y_1, \cdots, Y_k \leftarrow y_k]\varphi$ means that φ holds in the counterfactual world that would arise if Y_i is set to y_i, $i = 1, 2, \cdots, k$. We abbreviate such a formula as $[\vec{Y} \leftarrow \vec{y}]\varphi$.

A causal formula is a Boolean combination of basic causal formulas [1,32]. We use $(M, \vec{u}) \models \varphi$ to denote that a causal formula φ is true in causal model M given context $\vec{u}$. Given a recursive model M, $(M, \vec{u}) \models [\vec{Y} \leftarrow \vec{y}](X = x)$ if the value of X is x in the unique vector of values for the endogenous variables that simultaneously satisfies all equations $F_Z^{\vec{Y} \leftarrow \vec{y}}$, $Z \in \mathcal{V} - Y$ under the setting $\vec{u}$ of $\mathcal{U}$ [1,32]. Note that this definition can be extended to arbitrary causal formulas in the usual way.

Definition 2.1 *(Cause [32]).* We say that $\vec{X}=\vec{x}$ is a cause of φ in $(M,\vec{u})$ if the following three conditions hold:

AC1. $(M,\vec{u}) \models (\vec{X}=\vec{x}) \wedge \varphi$.
AC2. There exists a partition $(\vec{Z},\vec{W})$ of $\mathcal{V}$ with $\vec{X} \subseteq \vec{Z}$ and some setting $(\vec{x}',\vec{w}')$ of the variables in $(\vec{X},\vec{W})$ such that if $(M,\vec{u}) \models Z=z^*$ for $Z \in \vec{Z}$, then
 (a) $(M,\vec{u}) \models [\vec{X} \leftarrow \vec{x}', \vec{W} \leftarrow \vec{w}']\neg\varphi$. That is, changing $(\vec{X},\vec{W})$ from $(\vec{x},\vec{w})$ to $(\vec{x}',\vec{w}')$ changes φ from true to false.
 (b) $(M,\vec{u}) \models [\vec{X} \leftarrow \vec{x}, \vec{W} \leftarrow \vec{w}', \vec{Z}' \leftarrow \vec{z}^*]\varphi$ for all subsets $\vec{Z}'$ of $\vec{Z}-\vec{X}$. That is, setting $\vec{W}$ to $\vec{w}'$ should have no effect on φ as long as $\vec{X}$ has the value $\vec{x}$, even if all the variables in an arbitrary subset of $\vec{Z}$ are set to their original values in the context $\vec{u}$.
AC3. $(\vec{X}=\vec{x})$ is minimal, that is, no subset of $\vec{X}$ satisfies AC2.

As explained in [1], AC1 is used to capture the intuition that A cannot be a cause of B unless both A and B are true, while AC3 is a minimality condition ensuring that causes can always be taken to be single conjuncts. From now on, we are interested in $X=x$ rather than $\vec{X}=\vec{x}$. As the core of the definition of cause, AC2 emphasizes the important role of the contingency, which makes a distinction between the definition and the traditional counterfactual ones. In particular, if there is no variable in $\vec{W}$, then we call $\vec{X}=\vec{x}$ a *counterfactual cause* of φ in $(M,\vec{u})$ [31].

Definition 2.2 *(Degree of Responsibility [1]).* The degree of responsibility of $X=x$ for φ in $(M,\vec{u})$, denoted $dr((M,\vec{u}),(X=x),\varphi)$, is 0 if $X=x$ is not a cause of φ in $(M,\vec{u})$; it is $\frac{1}{k+1}$ if $X=x$ is a cause of φ in $(M,\vec{u})$ and there exists a partition $(\vec{Z},\vec{W})$ and setting x', $\vec{w}'$ for which AC2 holds such that (a) k variables in $\vec{W}$ have different values in $\vec{w}'$ than they do in the context $\vec{u}$ and (b) there is no partition $(\vec{Z}',\vec{W}')$ and setting x'', $\vec{w}''$ satisfying AC2 such that only $k'<k$ variables have different values in $\vec{w}''$ than they do in the context $\vec{u}$.

As stated in [1], the degree of responsibility of $X=x$ for φ in $(M,\vec{u})$ captures the minimal number of changes that have to be made in $\vec{u}$ in order to make φ counterfactually depend on X. In particular, if $X=x$ is a *counterfactual cause* of φ, then

$$dr((M,\vec{u}),(X=x),\varphi)=1.$$

We use the following example taken from [1] to illustrate the notion of degree of responsibility.

Example 2.1 *(Example 3.3 in [1]).* Consider a scenario that Mr. B wins an election against Mr. G by a vote of 11-0. Voter i is represented by a binary variable X_i, $i=1,2,\cdots,11$, which is 1 if voter i votes for Mr. B and 0 if voter i votes for Mr. G; the outcome is represented by the variable O, which is 1 if Mr. B wins and 0 if Mr. G wins. It is easy to check that each voter is a cause of Mr. B winning (that is, $X_i=1$ is a cause of $O=1$ for each i). The degree of responsibility of $X_i=1$ for $O=1$ is $\frac{1}{6}$, because at least five other voters must change their votes before changing X_i to 0 results in $O=0$.

3. The degree of responsibility of each formula for inconsistency

As mentioned earlier, the minimal inconsistent subsets of K may be considered as a characterization of inconsistency in K. From this point of view, it is natural to associate the problem of identifying the degree of responsibility of each formula of a knowledge base for inconsistency with the minimal inconsistent subsets of that base. Furthermore, the counterfactual dependence of the inconsistency in K on a given formula α of K can be expressed as that all the minimal inconsistent subsets of K would not have happened if α had not belonged to K under some contingency in intuition. Inspired by this, we propose a measurement for the degree of responsibility of each formula in a knowledge base for the inconsistency of that base in this section. We start with a notion of conditional minimal inconsistent subsets.

Definition 3.1. Let K be a knowledge base and α a formula of K. The set of conditional minimal inconsistent subsets of K given α, denoted $\mathsf{MI}(K|\alpha)$, is defined as

$$\mathsf{MI}(K|\alpha)=\{S \in \mathsf{MI}(K) | \alpha \notin S\}.$$

Evidently, the conditional minimal inconsistent subsets of K given α are exactly the minimal inconsistent subsets of K that cannot be broken by the removal of α from K. Moreover, $\mathsf{MI}(K|\alpha)=\mathsf{MI}(K)$ if α is a free formula of K, while $\mathsf{MI}(K|\alpha) \subset \mathsf{MI}(K)$ if $\alpha \in \bigcup \mathsf{MI}(K)$.

Definition 3.2. Let K be an inconsistent knowledge base and $\alpha \in \bigcup \mathsf{MI}(K)$. Then a (possibly empty) subset H of $\bigcup \mathsf{MI}(K|\alpha)$ is called a minimal hitting set of $\mathsf{MI}(K)$ given α, if

MH1. $\exists S \in \mathsf{MI}(K)$ such that $\alpha \in S$ and $H \cap S=\emptyset$;
MH2. $\forall S \in \mathsf{MI}(K|\alpha)$, $S \bigcap H \neq \emptyset$;
MH3. $|H|$ is minimal, that is, no H' with $|H'|<|H|$ satisfies MH1 and MH2.

Essentially, MH1 says that there is at least one minimal inconsistent subset containing α that cannot be broken by the removal of H from K. (Of course, such a minimal inconsistent subset can be broken by removing α from K.) MH2 says that all the conditional minimal inconsistent subsets of K given α can be broken if we delete all the formulas in H. Evidently, MH1 and MH2 together ensure that all the minimal inconsistent subsets of K can be broken if we delete $H \cup \{\alpha\}$ from K. In some sense, they capture the counterfactual dependence of the inconsistency in K on α in the case that H is removed from K. MH3 is a minimality condition to prevent more formulas from being involved in breaking minimal inconsistent subsets in $\mathsf{MI}(K|\alpha)$.

Evidently, if $\mathsf{MI}(K|\alpha) = \emptyset$, then $H = \emptyset$. In this special case, the inconsistency in K counterfactually depends on α, i.e., the inconsistency in K would not have happened if α had not belonged to K.

We use the following example to illustrate the two notions.

Example 3.1. Consider $K = \{a, b, \neg b \vee c, \neg c, \neg a, d, \neg d \wedge e, \neg e, f\}$. Then

$$K \vdash \bot \text{ and } \mathsf{MI}(K) = \{S_1, S_2, S_3, S_4\},$$

where

$$S_1 = \{a, \neg a\}, \quad S_2 = \{b, \neg b \vee c, \neg c\},$$
$$S_3 = \{d, \neg d \wedge e\}, \quad \text{and} \quad S_4 = \{\neg d \wedge e, \neg e\}.$$

Evidently, given $\neg c$, the set of conditional minimal inconsistent subsets of K is given as follows:

$$\mathsf{MI}(K|\neg c) = \{S_1, S_3, S_4\}.$$

It is easy to check that both

$$H_1 = \{a, \neg d \wedge e\}$$

and

$$H_2 = \{\neg a, \neg d \wedge e\}$$

are minimal hitting sets of $\mathsf{MI}(K)$ given $\neg c$.

Any minimal hitting set H of $\mathsf{MI}(K)$ given α must be one of the minimal correction subsets of $\bigcup \mathsf{MI}(K|\alpha)$, but not vice versa. On the other hand, $H \cup \{\alpha\}$ must be one of the smallest minimal correction subsets of K to which α belongs. But, allowing for MH1, H is not necessarily one of the smallest minimal correction subsets of $\bigcup \mathsf{MI}(K|\alpha)$. To illustrate this, consider the following example.

Example 3.2. Consider $K = \{a, \neg a \wedge b, \neg b, \neg b \wedge c\}$. Evidently,

$$\mathsf{MI}(K|a) = \{\{\neg a \wedge b, \neg b\}, \{\neg a \wedge b, \neg b \wedge c\}\}.$$

Then

$$H = \{\neg b, \neg b \wedge c\}$$

is the unique minimal hitting subset of $\mathsf{MI}(K)$ given a.

On the other hand, the singleton set $H' = \{\neg a \wedge b\}$ is the smallest minimal correction subset of $\bigcup \mathsf{MI}(K|a)$. But it does not satisfy *MH1*, because the removal of H' from K can break any minimal inconsistent subset of K.

Intuitively, a minimal hitting subset H of $\mathsf{MI}(K)$ given α is exactly one of the smallest sets of formulas that have to be removed together with α from K to break all the minimal inconsistent subsets of K under the constraint that at least one minimal inconsistent subset containing α cannot be broken by removing only H from K. Essentially, such a constraint ensures that α indeed plays a role in breaking all the minimal inconsistent subsets. From the viewpoint of the causal model presented in [32], H serves as a contingency in which the inconsistency of K counterfactually depends on α in some sense. So, we are more interested in the number of formulas in H.

Definition 3.3. Let K be an inconsistent knowledge base and $\alpha \in \bigcup \mathsf{MI}(K)$. Then the hitting value of K given α, denoted $f_h(K|\alpha)$, is defined as

$$f_h(K|\alpha) = |H|,$$

where H is a minimal hitting set of $\mathsf{MI}(K)$ given α.

Note that for any two minimal hitting sets H and H' of $\mathsf{MI}(K)$ given α, MH3 ensures that $|H| = |H'|$. Then $f_h(K|\alpha)$ associates α with a unique value. Essentially, the hitting value of K is the minimal number of formulas that have to be deleted to break all the minimal inconsistent subsets in $\mathsf{MI}(K|\alpha)$ under the constraint that at least one minimal inconsistent subset containing α remains unbroken.

Example 3.3. Consider $K = \{a, b, \neg b \vee c, \neg c, \neg a, d, \neg d \wedge e, \neg e, f\}$ again. Then

$$f_h(K|\neg c) = |H_1| = |H_2| = 2.$$

Now we are ready to define the degree of responsibility of each formula for the inconsistency of a knowledge base.

Definition 3.4. Let K be a knowledge base and α a formula of K. Then the degree of responsibility of α for the inconsistency of K, denoted $dr(K, \alpha)$, is defined as

$$dr(K,\alpha) = \begin{cases} 0, & \text{if } \alpha \in \mathsf{FREE}(K), \\ \frac{1}{1+f_h(K|\alpha)}, & \text{otherwise,} \end{cases}$$

where $f_h(K|\alpha)$ is the hitting value of K given α.

According to this definition, the degree of responsibility of any free formula for the inconsistency is 0. It coincides with the intuition that free formulas have nothing to do with minimal inconsistent subsets of K. Moreover, for $\alpha \in \bigcup \mathsf{MI}(K)$, $f_h(K|\alpha)$ gives the minimal number of formulas that need to be removed together with α from K in order to break all the minimal inconsistent sets of K. It grasps the minimal number of changes to obtain a contingency where the inconsistency counterfactually depends on α. In this sense, the measurement $dr(K, \alpha)$ captures the intuition of the notion of responsibility of A for B presented by Chockler and Halpern [1]. We will show that the measurement defined in such a way can be well explained in the causal model for inconsistency later. Here we use the following example to illustrate $dr(K, \alpha)$.

Example 3.4. Consider $K = \{a, b, \neg b \vee c, \neg c, \neg a, d, \neg d \wedge e, \neg e, f\}$ again. Then

$$dr(K,a) = dr(K,\neg a) = \frac{1}{3}, \quad dr(K, f) = 0.$$
$$dr(K,b) = dr(K,\neg b \vee c) = dr(K,\neg c) = \frac{1}{3},$$
$$dr(K,d) = dr(K,\neg e) = \frac{1}{4}, \quad dr(K,\neg d \wedge e) = \frac{1}{3}.$$

On the other hand, the following proposition shows that $dr(K, \alpha)$ can be also given in terms of minimal correction subsets of K. This accords with the duality of minimal inconsistent subsets and minimal correction subsets.

Proposition 3.1. *Let K be a knowledge base and $\alpha \in K$. Then*

$$dr(K,\alpha) = \begin{cases} \max\{\frac{1}{|R|} | R \in \mathsf{MC}(K)\ s.t.\ \alpha \in R\}, & \textit{if } \exists R \in \mathsf{MC}(K)\ s.t.\ \alpha \in R, \\ 0, & \textit{otherwise}. \end{cases}$$

Proof. Given a knowledge base K, $\alpha \in \bigcup \mathsf{MI}(K)$ if and only if there exists $R \in \mathsf{MC}(K)$ s.t. $\alpha \in R$. Without loss of generality, suppose that K is an inconsistent knowledge base. We only need to consider $\alpha \in \bigcup \mathsf{MI}(K)$. Let H be a minimal hitting set of $\mathsf{MI}(K)$ given α, then

$$H \cup \{\alpha\} \in \mathsf{MC}(K), \text{and } \alpha \in H \cup \{\alpha\}.$$

Moreover, MH3 guarantees that

$$|H \cup \{\alpha\}| = 1 + |H| \le |R|$$

for all $R \in \mathsf{MC}(K)$ s.t. $\alpha \in R$. Then

$$dr(K,\alpha) = \frac{1}{1+|H|} = \max\{\frac{1}{|R|} | R \in \mathsf{MC}(K)\ s.t.\ \alpha \in R\}.$$

So,

$$dr(K,\alpha) = \begin{cases} \max\{\frac{1}{|R|} | R \in \mathsf{MC}(K)\ s.t.\ \alpha \in R\}, & \text{if } \exists R \in \mathsf{MC}(K)\ s.t.\ \alpha \in R, \\ 0, & \text{otherwise.} \end{cases} \qquad \square$$

This proposition shows that we can measure the degree of responsibility of an individual formula $\alpha \in \bigcup \mathsf{MI}(K)$ by using the smallest minimal correction subsets of K that contain α. This provides an *alternative* expression of the degree of responsibility of each formula in K for the inconsistency of K in terms of minimal correction subsets.

Now we use the following example to illustrate this alternative expression.

Example 3.5. Consider $K = \{a, \neg a, a \wedge b, \neg b, c\}$. Then

$$\mathsf{MC}(K) = \{R_1, R_2, R_3\},$$

where $R_1 = \{a, a \wedge b\}$, $R_2 = \{\neg a, \neg b\}$, and $R_3 = \{\neg a, a \wedge b\}$. Then

$$dr(K, a) = \frac{1}{|R_1|} = \frac{1}{2}; \;\; dr(K, \neg b) = \frac{1}{|R_2|} = \frac{1}{2};$$

$$dr(K, \neg a) = \max\{\frac{1}{|R_2|}, \frac{1}{|R_3|}\} = \frac{1}{2};$$

$$dr(K, a \wedge b) = \max\{\frac{1}{|R_1|}, \frac{1}{|R_3|}\} = \frac{1}{2};$$

$$dr(K, c) = 0.$$

4. A causality-based explanation

In this section, we show that the measurement $dr(K, \alpha)$ captures exactly the degree of responsibility of α for the inconsistency of K if we formulate the inconsistency of K by using Halpern and Pearl's causal model.

To this end, we formulate the problem of inconsistency arising in a knowledge base by Halpern and Pearl's causal model [32] firstly. Given a knowledge base K and the set of minimal inconsistent subsets of K, to construct a causal model for the inconsistency of K,

- we associate every formula $\alpha \in K$ with a binary variable X_α, whose value is 1 if α keeps unchanged and 0 if α is deleted from K. We use $\vec{X}$ to denote the vector of all the variables corresponding to formulas.
- we associate with every minimal inconsistent subset $S \in \mathsf{MI}(K)$ a binary variable Y_S, whose value is 1 if S keeps unchanged and 0 if S is broken. We use $\vec{Y}$ to denote the vector of all the variables corresponding to minimal inconsistent subsets.
- the problem of inconsistency in K is represented by the binary variable I, whose value is 1 if K is inconsistent and 0 otherwise.

Let $\mathcal{V}_K = \{X_\alpha | \alpha \in K\} \cup \{Y_S | S \in \mathsf{MI}(K)\} \cup \{I\}$. Then $\mathcal{R}_K(V) = \{0, 1\}$ for each $V \in \mathcal{V}_K$.

Without loss of generality, we associate every variable X_α with a binary exogenous variable U_α. Moreover, we assume that the value of X_α depends on only the value of U_α. Let $\mathcal{U}_K = \{U_\alpha | \alpha \in K\}$. We use $\vec{U}$ to denote the vector of all the exogenous variables corresponding to formulas.

In addition, we use $\vec{X} - X_\alpha$ (resp. $\vec{Y} - Y_S$) to denote a vector that results from deleting X_α from $\vec{X}$ (resp. deleting Y_S from $\vec{Y}$).

We define the following functions:

- $F_{X_\alpha}(\vec{X} - X_\alpha, \vec{Y}, I, \vec{U}) = U_\alpha$ ($X_\alpha = U_\alpha$ for short) for every formula $\alpha \in K$.
- $F_{Y_S}(\vec{X}, \vec{Y} - Y_S, I, \vec{U}) = \prod_{\alpha \in S} X_\alpha$ ($Y_S = \prod_{\alpha \in S} X_\alpha$ for short) for every minimal inconsistent subset $S \in \mathsf{MI}(K)$.
- $F_I(\vec{X}, \vec{Y}, \vec{U}) = \bigoplus_{S \in \mathsf{MI}(K)} Y_S$ ($I = \bigoplus_{S \in \mathsf{MI}(K)} Y_S$ for short), where $\bigoplus$ is the Boolean addition.

Roughly speaking, the function F_{X_α} describes our assumption that the value of X_α depends on only the value of the exogenous variable U_α. In particular, $F_{X_\alpha}(U_\alpha) = 1$ if and only if $U_\alpha = 1$. Then the context $\vec{u} = (1, 1, \cdots, 1)$ ($\vec{u} = \vec{1}$ for short) describes the case that none of the formulas is deleted from K.

The function F_{Y_S} aims to capture the fact that we need to remove only one formula from the minimal inconsistent subset S to break S. The function F_I accords with the fact that we need to break all the minimal inconsistent subsets of K to restore consistency in K. In summary, the two kinds of functions capture the inherent features of inconsistency characterization in terms of minimal inconsistent subsets.

Now we are ready to construct a causal model for inconsistency. Let K be a knowledge base, then a causal model for the problem of inconsistency of K, denoted M_K, is defined as $M_K = \langle \mathcal{S}_K, \mathcal{F}_K \rangle$, where

$$\mathcal{S}_K = \langle \mathcal{U}_K, \mathcal{V}_K, \mathcal{R}_K \rangle$$

and

$$\mathcal{F}_K = \{F_{X_\alpha} | \alpha \in K\} \cup \{F_{Y_S} | S \in \mathsf{MI}(K)\} \cup \{F_I\}.$$

We use the following example to illustrate the notion of causal model for the inconsistency.

Example 4.1. Consider $K = \{a, b, \neg b \vee c, \neg c, \neg a, d, \neg d \wedge e, \neg e, f\}$ again. Then

$$\mathsf{MI}(K) = \{S_1, S_2, S_3, S_4\},$$

where

$$S_1 = \{a, \neg a\}, \quad S_2 = \{b, \neg b \vee c, \neg c\},$$
$$S_3 = \{d, \neg d \wedge e\}, \quad \text{and} \quad S_4 = \{\neg d \wedge e, \neg e\}.$$

Now we construct the causal model M_K as follows:

- Let U_α and X_α be the exogenous and endogenous binary variables corresponding to α for $\alpha \in K$, respectively. Then

  $$X_\alpha = U_\alpha \text{ for all } \alpha \in K.$$

- Let Y_{S_i} be the binary variable corresponding to S_i for $i = 1, 2, 3, 4$. Then

  $$Y_{S_1} = X_a \times X_{\neg a}, \quad Y_{S_2} = X_b \times X_{\neg b \vee c} \times X_{\neg c},$$
  $$Y_{S_3} = X_d \times X_{\neg d \wedge e}, \quad Y_{S_4} = X_{\neg d \wedge e} \times X_{\neg e}.$$

- Let I be the binary variable corresponding to the inconsistency of K, then

  $$I = Y_{S_1} \oplus Y_{S_2} \oplus Y_{S_3} \oplus Y_{S_4}.$$

Given a context $\vec{u} = \vec{1}$ (i.e., $U_\alpha = 1$ for all $\alpha \in K$), then

$$(M_K, \vec{u}) \models (X_\alpha = 1) \text{ for } \alpha \in K;$$
$$(M_K, \vec{u}) \models (Y_{S_i} = 1), \text{ for } i = 1, 2, 3, 4;$$
$$(M_K, \vec{u}) \models (I = 1).$$

Furthermore, consider the counterfactual world arising from $(X_a, X_b, X_d) \leftarrow \vec{0}$, then

$$(M_K, \vec{u}) \models [(X_a, X_b, X_d) \leftarrow \vec{0}](Y_{S_i} = 0), \text{ for } i = 1, 2, 3;$$
$$(M_K, \vec{u}) \models [(X_a, X_b, X_d) \leftarrow \vec{0}](Y_{S_4} = 1);$$
$$(M_K, \vec{u}) \models [(X_a, X_b, X_d) \leftarrow \vec{0}](I = 1).$$

The first two state that the first three minimal inconsistent subsets will be broken if we delete the formulas a, b, and d from K, whilst S_4 remains unbroken. The third states that the inconsistency would not disappear in the counterfactual world.

Consider another counterfactual world arising from $(X_a, X_b, X_{\neg d \wedge e}) \leftarrow \vec{0}$, then

$$(M_K, \vec{u}) \models [(X_a, X_b, X_{\neg d \wedge e}) \leftarrow \vec{0}](Y_{S_i} = 0), \text{ for } i = 1, 2, 3, 4;$$
$$(M_K, \vec{u}) \models [(X_a, X_b, X_{\neg d \wedge e}) \leftarrow \vec{0}](I = 0).$$

These coincide with the intuition that all the minimal inconsistent subsets will be broken if we delete the formulas a, b, and $\neg d \wedge e$ from K.

This causal model for the inconsistency of K can be also represented by the causal network illustrated in Fig. 1.

The following proposition provides an explanation for the measurement $dr(K, \alpha)$ from the point of view of causality.

Proposition 4.1. *Let K be a knowledge base and α a formula of K. Then*

1. *$X_\alpha = 1$ is a cause of $I = 1$ in $(M_K, \vec{1})$ if and only if $\alpha \in \bigcup \mathsf{MI}(K)$.*
2. *$dr((M_K, \vec{1}), (X_\alpha = 1), (I = 1)) = dr(K, \alpha)$.*

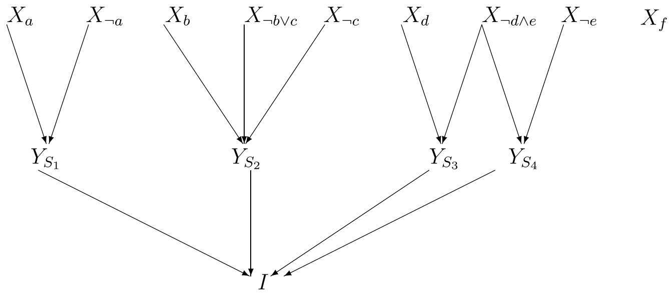

Fig. 1. The causal network for the inconsistency of K.

Proof. Let K be a knowledge base and $M_K = \langle \mathcal{S}_K, \mathcal{F}_K \rangle$ the causal model for the inconsistency of K.

1. $\Leftarrow$. If $\alpha \in \bigcup \mathsf{MI}(K)$, we only need to check AC2.
   Consider a minimal hitting set H of K given α. Let $\vec{W} = \vec{X}_H$, where $\vec{X}_H$ is the vector of variables corresponding to the formulas in H.
   - AC2 (a). If $X_\alpha = 0$, then $Y_T = 0$ for every T s.t. $\alpha \in T$. Moreover, if $\vec{W} = \vec{0}$, then $Y_{T'} = 0$ for every $T' \in \mathsf{MI}(K|\alpha)$. So, if $X_\alpha = 0$ and $\vec{W} = \vec{0}$, then $Y_S = 0$ for every $S \in \mathsf{MI}(K)$. This implies that $I = 0$. That is,
   $$(M_K, \vec{1}) \models [X_\alpha \leftarrow 0, \vec{W} \leftarrow \vec{0}]\neg(I = 1).$$
   - AC2 (b). Consider $\vec{Z} = \vec{X} - \vec{W}$. If $X_\alpha = 1$ and $\vec{Z}' = \vec{1}$ for all subsets $\vec{Z}'$ of $\vec{Z} - \vec{X}_\alpha$, then from MH1, there is at least one minimal inconsistent subset T s.t. $\alpha \in T$ and $Y_T = 1$ in the case of $\vec{W} = \vec{0}$. That is,
   $$(M_K, \vec{1}) \models [X_\alpha \leftarrow 1, \vec{W} \leftarrow \vec{0}, \vec{Z}' = \vec{1}](I = 1).$$

   So, $X_\alpha = 1$ is a cause of $I = 1$.
   $\Rightarrow$. Suppose that $X_\alpha = 1$ is a cause of $I = 1$. If $\alpha \notin \bigcup \mathsf{MI}(K)$, then we cannot find a partition $(\vec{Z}, \vec{W})$ satisfying AC2, because the value of I has nothing to do with the values of the variables corresponding to free formulas in M_K. Then $\alpha \in \bigcup \mathsf{MI}(K)$.
2. From the definition of the minimal hitting set given α, we know that no subset of $\vec{W} = \vec{X}_H$ satisfies AC2. Then
$$\begin{aligned} dr((M_K, \vec{1}), (X_\alpha = 1), (I = 1)) &= \begin{cases} 0, & \text{if } \alpha \in \mathsf{FREE}(K) \\ \frac{1}{1 + f_h(K|\alpha)}, & \text{otherwise} \end{cases} \\ &= dr(K, \alpha). \qquad \square \end{aligned}$$

The first item of this proposition shows that only formulas involved in minimal inconsistent subsets may be considered as causes of the inconsistency in a knowledge base. This accords with that only any formula involved in minimal inconsistent subsets has to bear a nonzero responsibility for the inconsistency of a knowledge base in the context of the measurement $dr(K, \alpha)$. The second item shows that the measurement $dr(K, \alpha)$ exactly grasps the degree of responsibility of α for the inconsistency of K from the point of view of causality.

5. Logical properties and computational complexity

In this section, some interesting behaviors of the measurement dr are illustrated by several properties. The computational complexity issue about the measurement is discussed as well.

The measurement dr uses a value in $[0, 1]$ to grasp the responsibility of a given formula for the inconsistency in a knowledge base from the point of view of causality. The following proposition shows that the special values 0 and 1 correspond to free formulas and common formulas of all the minimal inconsistent subsets, respectively.

Proposition 5.1. *Let K be a knowledge base and $\alpha \in K$. Then*

(1) $dr(K, \alpha) = 0$ *if and only if α is a free formula of K.*
(2) $dr(K, \alpha) = 1$ *if and only if $\alpha \in \bigcap \mathsf{MI}(K)$.*

Proof. Obviously, (1) is a direct consequence of the definition of dr.

We consider (2) only. Let K be a knowledge base and $\alpha \in K$, then

$$dr(K, \alpha) = 1 \Leftrightarrow f_h(K|\alpha) = 0 \Leftrightarrow \mathsf{MI}(K|\alpha) = \emptyset \Leftrightarrow \alpha \in \bigcap \mathsf{MI}(K). \qquad \square$$

Note that the property (1) is termed as *Minimality* in [16,20]. Clearly, *Minimality* states that free formulas bear no responsibility for the inconsistency. This coincides with the intuition that any free formula is not a cause of the inconsistency. In contrast, $dr(K,\alpha) = 1$ means that α can take full responsibility for the inconsistency, i.e., α is a counterfactual cause of the inconsistency. In this sense, the property (2) shows that common formulas of all the minimal inconsistent subsets are exactly the counterfactual causes of the inconsistency. Intuitively, this coincides with the case that all such formulas are involved in each of the minimal inconsistent subsets of K.

Corollary 5.1. *Let K be a minimal inconsistent knowledge base. Then*

$$\forall \alpha, \beta \in K,\ dr(K,\alpha) = dr(K,\beta) = 1.$$

Proof. This is a direct consequence of Proposition 5.1, because $K = \bigcap \mathsf{MI}(K)$ if K is a minimal inconsistent knowledge base. □

This corollary has dual aspects. One states that every formula of a minimal inconsistent knowledge base is a counterfactual cause of the inconsistency in that base. Another states that any two formulas of a minimal inconsistent knowledge base have the same degree of responsibility for the inconsistency of that base, which is termed as the property of *Fairness* in [21].

The following proposition shows that the special value $\frac{1}{|K|}$ corresponds to the case that K consists of self-contradictory formulas.

Proposition 5.2. *Let K be a knowledge base and $\alpha \in K$. Then*

$$dr(K,\alpha) = \frac{1}{|K|}$$

if and only if each formula of K is inconsistent.

Proof. $dr(K,\alpha) = \frac{1}{|K|} \Longleftrightarrow \mathsf{MC}(K) = \{K\} \Longleftrightarrow \forall \alpha \in K,\ \alpha \vdash \bot.$ □

The following proposition shows that adding a free formula to a knowledge base cannot affect the degree of responsibility of each formula for the inconsistency of that base. It coincides with the intuition that free formulas bring no new minimal inconsistent subset. Generally, such a property is termed as *Free Formula Independence* in [16,20].

Proposition 5.3. *Let K be a knowledge base and $\alpha \in K$. If β is a free formula of $K \cup \{\beta\}$, then $dr(K,\alpha) = dr(K \cup \{\beta\}, \alpha)$.*

Proof. Consider the alternative expression of $dr(K,\alpha)$. If β is a free formula, then

$$\mathsf{MC}(K \cup \{\beta\}) = \mathsf{MC}(K).$$

So, for every formula $\alpha \in K$,

$$dr(K,\alpha) = dr(K \cup \{\beta\}, \alpha). \qquad \square$$

On the other hand, the following example shows that adding a formula to a knowledge base may make different impacts on the degrees of responsibility of different formulas in that base if such an extension brings some new minimal inconsistent subsets.

Example 5.1. Consider $K = \{a, \neg a, b\}$. Evidently,

$$dr(K,a) = dr(K,\neg a) = 1,\ dr(K,b) = 0.$$

If we add $\neg b$ to K, then

$$dr(K \cup \{\neg b\}, a) = \frac{1}{2} < dr(K,a),\ dr(K \cup \{\neg b\}, \neg a) = \frac{1}{2} < dr(K,\neg a).$$

But

$$dr(K \cup \{\neg b\}, b) = \frac{1}{2}\ >\ dr(K,b).$$

Actually, the following proposition shows how responsibilities of formulas change when adding a formula to that base.

Proposition 5.4. *Let K be a knowledge base and $\alpha \in K$. Let β be a new formula not in K.*

(1) *If $dr(K, \alpha) > 0$, then $dr(K, \alpha) \geq dr(K \cup \{\beta\}, \alpha)$.*
(2) *If $dr(K, \alpha) = 0$, then $dr(K, \alpha) \leq dr(K \cup \{\beta\}, \alpha)$.*

Proof. Given a knowledge base K and $\alpha \in K$.

(1) If $dr(K, \alpha) > 0$, consider the alternative expression of $dr(K, \alpha)$. If β is a new formula, then

$$\mathsf{MI}(K|\alpha) \subseteq \mathsf{MI}(K \cup \{\beta\}|\alpha).$$

Let H' and H be minimal hitting sets of K and $K \cup \{\beta\}$ given α, respectively, then
- $|H'| \leq |H|$ if $\beta \notin H$.
- $|H'| < |H|$ if $\beta \in H$, because $|H'| \leq |H \setminus \{\beta\}|$.

So,

$$dr(K, \alpha) \geq dr(K \cup \{\beta\}, \alpha).$$

(2) $dr(K \cup \{\beta\}, \alpha) \geq 0 = dr(K, \alpha)$. □

Here (1) says that the degree of responsibility of α may decrease by adding a formula to K if α is a cause of the inconsistency in K. This coincides with the intuition that more formulas may be involved in the inconsistency in $K \cup \{\beta\}$. In contrast, (2) states that a free formula of K may be involved in new inconsistency caused by adding a formula to K.

Recall that the counterfactual dependence under some contingency plays a crucial role in identifying the degree of responsibility of each formula of a knowledge base for the inconsistency of that base. Essentially, given a cause α of the inconsistency, $dr(K, \alpha)$ depends on the minimal number of other potential causes of the inconsistency. This leads to the following property.

Proposition 5.5. *Let K and K' be two knowledge bases and $\alpha \in \bigcup \mathsf{MI}(K)$. If $f_h(K|\alpha) = f_h(K \cup K'|\alpha)$, then $dr(K, \alpha) = dr(K \cup K', \alpha)$.*

Proof. This is a direct consequence of Definition 3.4. □

We call this property *Hitting Value Invariance*. It states that an extension of K cannot affect the responsibility of a cause for the inconsistency if it does not change the hitting value given that cause. From the point of view of causality, this property grasps the essence that the degree of responsibility of a formula for the inconsistency is determined by the minimal number of changes that need to obtain a contingency where the inconsistency counterfactually depends on the formula.

Corollary 5.2. *Let K and K' be two knowledge bases and $\alpha \in \bigcup \mathsf{MI}(K)$. If $\mathsf{MI}(K|\alpha) = \mathsf{MI}(K \cup K'|\alpha)$, then $dr(K, \alpha) = dr(K \cup K', \alpha)$.*

Proof. Given $\alpha \in \bigcup \mathsf{MI}(K)$, if $\mathsf{MI}(K|\alpha) = \mathsf{MI}(K \cup K'|\alpha)$, then $f_h(K|\alpha) = f_h(K \cup K'|\alpha)$. From Proposition 5.5, $dr(K, \alpha) = dr(K \cup K', \alpha)$. □

We call this property *α-Conditional Independence*. It states that an extension of K cannot affect the responsibility of a cause α for the inconsistency if it does not bring any new conditional minimal inconsistent subset given α.

Both *Hitting Value Invariance* and *α-Conditional Independence* are different from *Free Formula Independence*. The former two properties describe conditions such that the responsibility of *a given cause* of the inconsistency in the original knowledge base remains unchanged when we extend the base. In contrast, *Free Formula Independence* describes a condition such that the responsibility of *each formula* in the original knowledge base for the inconsistency remains unchanged when we add a new formula to the base. On the other hand, note that $\mathsf{MI}(K|\alpha) = \mathsf{MI}(K \cup \{\beta\}|\alpha)$ for any free formula β if $\alpha \in \bigcup \mathsf{MI}(K)$. Then the three properties coincide with each other in characterizing the impact of adding free formulas to a knowledge base on the responsibility of each cause of the inconsistency in the original knowledge base.

Both *Hitting Value Invariance* and *α-conditional Independence* focus on unchanged responsibility for the inconsistency. The following proposition shows a special case that we can compute the responsibility of a cause for the inconsistency in the combination of two knowledge bases from the respective responsibilities of that cause for the inconsistencies in the two knowledge bases.

Proposition 5.6. *Let K and K' be two knowledge bases. Let $\alpha \in (\bigcup \mathsf{MI}(K)) \cap (\bigcup \mathsf{MI}(K'))$. If $\forall M \in \mathsf{MI}(K|\alpha), \forall M' \in \mathsf{MI}(K'|\alpha)$, $M \cap M' = \emptyset$, then*

$$\frac{1}{dr(K \cup K', \alpha)} = \frac{1}{dr(K, \alpha)} + \frac{1}{dr(K', \alpha)} - 1.$$

Proof. Consider $\alpha \in (\bigcup \mathsf{MI}(K)) \cap (\bigcup \mathsf{MI}(K'))$. If $\forall M \in \mathsf{MI}(K|\alpha), \forall M' \in \mathsf{MI}(K'|\alpha)$, $M \cap M' = \emptyset$, then $H \cap H' = \emptyset$ for any two minimal hitting sets H and H' of K and K' given α, respectively. Moreover, $H \cup H'$ must be a minimal hitting set of $K \cup K'$ given α. So,

$$\frac{1}{dr(K \cup K', \alpha)} = |H| + |H'| + 1 = \frac{1}{dr(K, \alpha)} + \frac{1}{dr(K', \alpha)} - 1. \qquad \square$$

Given two knowledge bases and a common cause of the inconsistencies in the two bases, this property states that the responsibility of that common cause for the inconsistency in the combination of two bases can be explicitly represented by the responsibility of that cause for the inconsistency in each base if any two conditional minimal inconsistent subsets of different knowledge bases do not overlap each other. Moreover, such a representation can be explained as follows: given a common cause for the inconsistencies in two knowledge bases, if the other potential causes for the respective inconsistencies in two knowledge bases are completely different, then the combination of the other potential causes for the respective inconsistencies are exactly the other potential causes for the inconsistency in the combination of two knowledge bases.

Lastly, the following example shows that replacing a formula with another logically stronger formula has different impacts on the degrees of responsibility of different formulas.

Example 5.2. Consider $K = \{a, \neg a, \neg b\}$ and $K' = (K \setminus \{a\}) \cup \{a \wedge b\}$. Then $dr(K, a) = dr(K', a \wedge b) = 1$, $dr(K, \neg a) = 1 > \frac{1}{2} = dr(K', \neg a)$, but $dr(K, \neg b) = 0 < \frac{1}{2} = dr(K', \neg b)$.

Now we turn to complexity issue. We assume that the reader is familiar with the basics of complexity, in particular the polynomial hierarchy.

It has been shown that computing the degree of responsibility in binary models is $FP^{NP[\log n]}$-complete [36]. In general case, it has been shown that computing the degree of responsibility is $FP^{\Sigma_2^P[\log n]}$-complete in general recursive models [1].

At first, we give the following proposition presented in [36].

Proposition 5.7. *Computing the degree of responsibility is $FP^{NP[\log n]}$-complete in binary causal models.*

Then we can get the following corollary.

Corollary 5.3. *Computing the degree of responsibility of each formula in a knowledge base for the inconsistency is $FP^{NP[\log n]}$-complete in the case of the set of minimal inconsistent subsets is given.*

Proof. Given a knowledge base K and the set $\mathsf{MI}(K)$ of minimal inconsistent subsets of K, it is easy to check that we can construct a binary causal model M_K for K in polynomial time, by following the procedure mentioned in Section 4. Then according to Proposition 5.7, computing the degree of responsibility of each formula in a knowledge base for the inconsistency is $FP^{NP[\log n]}$-complete. $\square$

Recall that the measurement $dr(K, \alpha)$ can be also given in terms of minimal correction subsets. Given a knowledge base K and a formula $\alpha \in K$, we have the following observations:

- $dr(K, \alpha) \in \{0, 1, \frac{1}{2}, \cdots, \frac{1}{i}, \cdots, \frac{1}{|K|}\}$.
- If K is consistent, then $dr(K, \alpha) = 0$.
- $dr(K, \alpha) = 1$ if and only if $K \vdash \bot$ but $K \setminus \{\alpha\} \nvdash \bot$.
- If $K \vdash \bot$ and $K \setminus \{\alpha\} \vdash \bot$, then $0 \le dr(K, \alpha) \le \frac{1}{2}$.
- If $K \vdash \bot$ and $dr(K, \alpha) < \frac{1}{|K|}$, then $dr(K, \alpha) = 0$.

These observations can help us identify the following complexity result about computing $dr(K, \alpha)$.

Proposition 5.8. *Let K be a knowledge base and α a formula of K, computing $dr(K, \alpha)$ is in $FP^{\Sigma_2^P[\log n]}$.*

Proof. As we have observed, the number of possible values of $dr(K, \alpha)$ is $|K| + 1$. Moreover, for two special values 0 and 1, we have the following results, respectively:

- $dr(K, \alpha) = 1$ if and only if $K \vdash \bot$ but $K \setminus \{\alpha\} \nvdash \bot$.
- $dr(K, \alpha) = 0$ if $K \nvdash \bot$, or $K \vdash \bot$ and $dr(K, \alpha) < \frac{1}{|K|}$.

Furthermore, consider the following nondeterministic algorithm for deciding whether $dr(K,\alpha) \geq \frac{1}{i}$ given $i \in \{2, \cdots, |K|\}$:

- guess a subset H of $K \setminus \{\alpha\}$ that contains at most $i-1$ formulas;
  (a) check that $K \setminus H \vdash \perp$;
  (b) **if** so, check that $K \setminus (H \cup \{\alpha\}) \nvdash \perp$;
  (c) **if** so, check the minimality of H;
  (d) **if** so, $dr(K,\alpha) \geq \frac{1}{i}$;

Note that checking whether a knowledge base is consistent is in *NP*, and checking whether a knowledge base is inconsistent is in *coNP*. So, this query is in Σ_2^P.

Lastly, let us consider the following algorithm for computing $dr(K,\alpha)$ for a given K and a given formula $\alpha \in K$:

1. **if** $K \nvdash \perp$, **then** $dr(K,\alpha) = 0$;
2. **else if** $K \setminus \{\alpha\} \nvdash \perp$, **then** $dr(K,\alpha) = 1$;
3. **else** check that $dr(K,\alpha) \geq \frac{1}{|K|}$;
   (a) **if** not, $dr(K,\alpha) = 0$;
   (b) **else** performs a binary search on $\{\frac{1}{2}, \cdots, \frac{1}{|K|-1}\}$ for $dr(K,\alpha)$, with a call to the Σ_2^P oracle above to check whether $dr(K,\alpha) \geq \frac{1}{i}$.

According to this algorithm, in the case that $K \setminus \{\alpha\} \vdash \perp$, $dr(K,\alpha)$ can be computed using $O(\log|K|)$ calls to the Σ_2^P oracle. So, computing $dr(K,\alpha)$ is in $FP^{\Sigma_2^P[\log n]}$. $\square$

6. Related work

In this section, we compare our measurement *dr* with some closely related measurements.

To the best of our knowledge, the Shapley inconsistency value presented by Hunter and Konieczny [5,20,16] is the first attempt to capture the contribution/responsibility of an individual formula in a knowledge base for the inconsistency in that base. Roughly speaking, the Shapley inconsistency value uses the Shapley value, one of the well known cooperation game models, to distribute an assessment for the overall inconsistency in a knowledge base to each formula in that base. Then the part assigned to an individual formula, termed the Shapley inconsistency value of the formula, may be considered as a measurement for the contribution/responsibility of that formula to the inconsistency [5,20,16]. To be more precise, given a knowledge base K and a formula $\alpha \in K$, the Shapley inconsistency value of α, denoted $S_\alpha^I(K)$, is given as

$$S_\alpha^I(K) = \sum_{C \subseteq K} \frac{(|C|-1)!(|K|-|C|)!}{|K|!}(I(C) - I(C \setminus \{\alpha\})),$$

where I is an inconsistency measure for assessing the overall inconsistency in any subset of K [5,20,16]. In particular, in the case that $I(K) = I_{MI}(K) = |\mathsf{MI}(K)|$, the Shapley inconsistency value can be defined alternatively as follows [20,16]:

$$S_\alpha^{I_{MI}}(K) = MIV_C(K,\alpha) = \sum_{M \in \mathsf{MI}(K)\ s.t.\ \alpha \in M} \frac{1}{|M|}.$$

In addition, Mu et al. have proposed an opposed formula-based approach for identifying the blame of each formula for the inconsistency in a prioritized knowledge base [37]. However, it has been shown that the blame of a formula α for the inconsistency in K defined in [37] is exactly $S_\alpha^{I_{MI}}(K)$ (i.e., $MIV_C(K,\alpha)$) in the case that K is a classical (flat or non-prioritized) knowledge base [37]. Allowing for this, here we compare $dr(K,\alpha)$ with $S_\alpha^I(K)$ in general and $S_\alpha^{I_{MI}}(K)$ in particular, respectively.

The measurement *dr* differs from the Shapley inconsistency value in the following aspects.

- At first, their respective starting points are different from each other. The Shapley inconsistency value aims to capture the *contribution* made by each formula in a knowledge base to the inconsistency of that base by using cooperation game models. The model of Shapley value plays a crucial role in identifying the contribution of each formula to the inconsistency. In contrast, the goal of the measurement *dr* is to grasp *the degree of responsibility* of each formula in a knowledge base for the inconsistency in that base from the point of view of causality. The notion of counterfactual dependence under some contingency is central to the degree of responsibility.
- Secondly, $S_\alpha^I(K)$ depends on the inconsistency measure I. It is a part of $I(K)$ distributed to α according to the Shapley value model if we consider any subset C of K and $I(C)$ as a coalitional game and its utility, respectively. Moreover, the behavior of the Shapley inconsistency value may depend on that of I. In contrast, identification of $dr(K,\alpha)$ is independent to any assessment for the overall inconsistency of a knowledge base. Instead, it depends on the minimal number of formulas that have to be removed from K in order to make the inconsistency in K counterfactually depend on α.

- Thirdly, $S^I_\alpha(K)$ only focuses on the part of inconsistency brought by α. As the marginal utility of α in a coalitional game, $S^I_\alpha(K)$ is a weighted accumulation of $I(C) - I(C \setminus \{\alpha\})$ for all $C \subseteq K$ s.t. $\alpha \in C$. Essentially, $I(C) - I(C \setminus \{\alpha\})$ assesses the part of inconsistency that α brings to C if $\alpha \in C$. In this sense, $S^I_\alpha(K)$ has nothing to do with the other part of the inconsistency that α is not involved in. In contrast, $dr(K,\alpha)$ is also interested in the other parts of inconsistency that α is not involved in if α is identified as a cause of the inconsistency. Informally speaking, it is interested in identifying the number of other potential causes of the inconsistency. However, as a particular Shapley inconsistency value, $S^{I_{MI}}_\alpha(K)$ illustrates this difference more clearly. Although $S^{I_{MI}}_\alpha(K)$ and $dr(K,\alpha)$ are given in terms of minimal inconsistent subsets, they focus on different parts of $\mathsf{MI}(K)$ if α is involved in the inconsistency. Given $\alpha \in \bigcup \mathsf{MI}(K)$, $S^{I_{MI}}_\alpha(K)$ is interested in only minimal inconsistent subsets that counterfactually depend on α, i.e., $\mathsf{MI}(K) \setminus \mathsf{MI}(K|\alpha)$. In contrast, $dr(K,\alpha)$ is more interested in conditional minimal inconsistent subsets given α, i.e., $\mathsf{MI}(K|\alpha)$.
- Lastly, as shown in [5], $\max\limits_{\alpha \in K} S^I_\alpha(K)$ can be considered as a measurement for the inconsistency in the whole knowledge base. However, $\max\limits_{\alpha \in K} dr(K,\alpha)$ is not suitable for measuring the inconsistency of K. To illustrate this, consider $K_1 = \{a, \neg a, b, c, d\}$ and $K_2 = \{a, \neg a, b, \neg b, d\}$. Note that $\max\limits_{\alpha \in K_1} dr(K_1,\alpha) = 1 > \frac{1}{2} = \max\limits_{\alpha \in K_2} dr(K_2,\alpha)$. But K_2 is more inconsistent that K_1 in intuition.

Intuitively, the degree of responsibility of a formula for the inconsistency should be related to the contribution made by the formula to the inconsistency. Actually, $S^{I_{MI}}_\alpha(K)$ and $dr(K,\alpha)$ have the following two aspects as common ground.

- At first, $S^{I_{MI}}_\alpha(K)$ coincides with $dr(K,\alpha)$ in the role of free formulas in causing the inconsistency. If α is a free formula in K, then $S^{I_{MI}}_\alpha(K) = 0$. This property implies that free formulas make no contribution to the inconsistency in K. It is consistent with that free formulas need not bear any responsibility for the inconsistency of K in the context of dr.
- Secondly, $S^{I_{MI}}_\alpha(K)$ accords with $dr(K,\alpha)$ in characterizing the indistinguishableness of formulas in a minimal inconsistent knowledge base. If K is a minimally inconsistent knowledge base, then $S^{I_{MI}}_\alpha(K) = \frac{1}{|K|}$ for each $\alpha \in K$. This property implies that all the formulas in K equally share the inconsistency in K. It accords with that any two formulas have the same degree of responsibility for the inconsistency in K in the context of dr (see Corollary 5.1).

However, $S^{I_{MI}}_\alpha(K)$ is not always consistent with $dr(K,\alpha)$. For example, consider $K = \{a, \neg a, b, \neg b, \neg b \vee c, \neg c\}$. Evidently, we can obtain the following results:

$$S^{I_{MI}}_a(K) = \frac{1}{2} < S^{I_{MI}}_b(K) = \frac{5}{6}, \text{ but } dr(K,a) = dr(K,b) = \frac{1}{2}.$$

This disaccord can be attributed to their different underlying ideas. In other words, their different starting points make the two measurements behave differently. The measurement dr is stemmed from causality. The counterfactual dependence under some contingency is at the very core of the notion of responsibility. As mentioned above, such a fundamental idea is partially characterized by the properties of *Hitting Value Invariance* and *α-Conditional Independence*. However, $S^{I_{MI}}_\alpha(K)$ does not satisfy the properties. On the other hand, $S^{I_{MI}}_\alpha(K)$ sums up contributions made by α to all the minimal inconsistent subsets, and it can be fully characterized by five axioms [16]. However, the following example shows that $dr(K,\alpha)$ does not accord with the axiom of *Decomposition*, one of the five axioms.

Example 6.1. Consider $K = \{a, \neg a, d, \neg d\}$ and $K' = \{a, \neg a \vee c, \neg c, b, \neg b\}$. Then

$$dr(K,a) = \frac{1}{2},$$
$$dr(K',a) = \frac{1}{2}.$$

Note that

$$|\mathsf{MI}(K \cup K')| = |\mathsf{MI}(K)| + |\mathsf{MI}(K')|,$$

but

$$dr(K \cup K',a) = \frac{1}{3} \neq dr(K,a) + dr(K',a).$$

Instead,

$$\frac{1}{dr(K \cup K',a)} = \frac{1}{dr(K,a)} + \frac{1}{dr(K',a)} - 1.$$

Besides the special inconsistency value $S_{\alpha}^{I_{MI}}$, Hunter et al. presented two other MinInc inconsistency values based on minimal inconsistent subsets in [20], namely MIV_D and $MIV_{\sharp}$. Let K be a knowledge base and $\alpha \in K$, then

$$MIV_D(K,\alpha) = \begin{cases} 1, & \text{if } \alpha \in \bigcup \mathsf{MI}(K), \\ 0, & \text{otherwise}, \end{cases} \quad \text{and}$$
$$MIV_{\sharp}(K,\alpha) = |\{S \in \mathsf{MI}(K) | \alpha \in S\}|.$$

Compared to dr, MIV_D is too simple to distinguish any two formulas involved in the inconsistency from each other in the degree of responsibility for the inconsistency. On the other hand, as the naive version of $S_{\alpha}^{I_{MI}}$, the MinInc inconsistency value $MIV_{\sharp}(K,\alpha)$ also depends on only the minimal inconsistent subsets that counterfactually depend on α. This makes both the difference and the common ground of dr and $S_{\alpha}^{I_{MI}}$ mentioned above also meaningful when we compare $MIV_{\sharp}$ and dr.

The *DIM* measures presented in [22] may be considered as an attempt to measure the degree of responsibility of each formula for the inconsistency in a knowledge base based on MUS-graph of that base. Roughly speaking, the MUS-graph of a knowledge base is a graph with nodes corresponding to minimal inconsistent subsets of that base and an edge between two nodes if the two corresponding minimal inconsistent subsets overlap each other. Then the distance $d_{MUS}(\alpha, S)$ between a formula S and a minimal inconsistent subset S is defined based on paths in the MUS-graph. In particular, $d_{MUS}(\alpha, S) = 0$ iff $\alpha \in S$. Finally, the *DIM* measures are given in terms of the max distance between a given formula and minimal inconsistent subsets. Allowing for the tight relation between the smallest minimal correction subsets and the overlaps of minimal inconsistent subsets, both the *DIM* measure and dr takes into account the impact of the structure of minimal inconsistent subsets on identifying the degree of each formula for the inconsistency. Moreover, $\mathsf{MI}(K|\alpha)$ is involved in both the *DIM* measure and dr. However, the way of $dr(K,\alpha)$ depending on $\mathsf{MI}(K|\alpha)$ can be well explained by the counterfactual dependence under some contingency from the perspective of causality. In contrast, the way of involving $\mathsf{MI}(K|\alpha)$ in the *DIM* measures lacks an intuitive explanation from causation of the inconsistency.

The measure $I_{\mathcal{P}_m}$ for the degree of responsibility of a formula for the inconsistency of a knowledge base presented in [25] aims to capture the contribution made by the formula to the inconsistency in that base based on minimal proofs. Roughly speaking, $I_{\mathcal{P}_m}(\alpha)$ sums up the times of the formula α involved in both minimal proofs of x and that of $\neg x$ for any variable x occurring in formulas of K. Such a fine-grained measure may allow some formulas involved in the inconsistency to bear no responsibility for the inconsistency. This makes it different from dr. To illustrate this, consider $K = \{a, \neg a, b, \neg b\}$, then $I_{\mathcal{P}_m}(a) = I_{\mathcal{P}_m}(b) = I_{\mathcal{P}_m}(\neg a) = I_{\mathcal{P}_m}(\neg b) = 0$. This implies that none of formulas bears responsibility for the inconsistency of K.

Although the process of identifying $dr(K,\alpha)$ need not depend on any inconsistency measure for K, there exists an interesting relation between $dr(K,\alpha)$ and some measure for quantifying the overall inconsistency of K. Let K be an inconsistent knowledge base and $\alpha \in K$. Suppose that $dr(K,\alpha) = \frac{1}{m}$, then there exists a minimal correction subset R such that $\alpha \in R$ and $|R| = m = \frac{1}{dr(K,\alpha)}$. However, the smallest size of minimal correction subsets of a knowledge base may be considered as an evaluation of effort to resolve the inconsistency in that base, and then can be used to measure the inconsistency in that base [18]. Here we give a simple version I_{dr} of the measure defined for syntax-based information base in [18] as follows: let K be a knowledge base, then

$$I_{dr}(K) = \min_{R \in \mathsf{MC}(K)} |R|.$$

Evidently, this measure can be expressed in terms of dr alternatively:

$$I_{dr}(K) = \begin{cases} \dfrac{1}{\max_{\alpha \in K} dr(K,\alpha)}, & \text{if } \exists \alpha \in K \text{ s.t. } dr(K,\alpha) > 0, \\ 0, & \text{else}. \end{cases}$$

On the other hand, the measurement dr provides a good clue for making a decision on resolving the inconsistency in a knowledge base from the point of view of causality. Intuitively, we should give priority to formulas with higher responsibility for the inconsistency rather than ones with lower responsibility when we have to change some formulas to resolve the inconsistency in general case. In this sense, the measurement dr may help us identify formulas to be changed. In addition, given an inconsistent knowledge base K, $I_{dr}(K)$ gives the minimal size of minimal correction subsets of K. Moreover, for any formula β involved in the inconsistency of K, there is a minimal correction subset R such that $\beta \in K$ and

$$|R| \leq \frac{1}{\min\limits_{\alpha \in K \text{ s.t. } dr(K,\alpha) > 0} dr(K,\alpha)}.$$

This may help us check whether some constraints on the number of formulas to be changed is satisfiable. Moreover, given a formula α involved in the inconsistency in K, $dr(K,\alpha)$ tells us that we need to remove only $\frac{1}{dr(K,\alpha)} - 1$ formulas together with α from K to restore consistency in K. This may help us check whether a proposal of inconsistency handling involves unnecessary formulas in the case that α must be involved in the proposal. We illustrate these aspects by the following example.

Example 6.2. Consider $K = \{a, \neg a, b, \neg b, \neg b \vee c, \neg c, d\}$. Then

$$dr(K, a) = dr(K, \neg a) = dr(K, b) = \frac{1}{2},\ dr(K, d) = 0.$$

$$dr(K, \neg b) = dr(K, \neg b \vee c) = dr(K, c) = \frac{1}{3}.$$

From the measurement dr, we can obtain the following observations about inconsistency resolving:

- To restore consistency in K, we have to remove at least 2 formulas from K, because $\max_{\alpha \in K} dr(K, \alpha) = \frac{1}{2}$.
- Given a formula α in K, we can find a minimal correction subset R such that $\alpha \in R$ and $2 \le |R| \le 3$, because $\max_{\alpha \in K \setminus \{d\}} dr(K, \alpha) = \frac{1}{3}$.

Lastly, the measurement dr can be given in terms of minimal correction subsets of a knowledge base alternatively. This implies that the measurement dr can be generalized from the propositional logic to some more complex logics in which the notion of minimal correction subset is exactly the same as that in propositional logic.

7. An application in requirements engineering

In this section we use a small but explanatory example in requirements engineering to illustrate the application of the measurement for the degree of responsibility for the inconsistency. We consider a scenario for eliciting requirements for updating an existing software, which is slightly adapted from the example used in [24,27].

Example 7.1. Consider the following scenario for eliciting requirements for updating an existing software. There are three stakeholders involved in this scenario, including the seller of the new system, the user of the existing system (the user for short), and the domain expert in requirements engineering. Each of the three stakeholders may provide demands from her/his own perspective. When inconsistencies in their demands are identified, developers and the three stakeholders start to negotiate on resolving inconsistencies. The measurement dr may help developers and the stakeholders make a decision on revising the requirements.

- The seller of the new system provides two demands:
  (a1) The user interface of the system-to-be should be in the modern idiom (i.e., fashionable).
  (a2) The system-to-be should be open, that is, the system-to-be could be extended easily.
- The user of the existing system provides three demands:
  (b1) The system-to-be should be developed based on the techniques used in the existing system.
  (b2) The user interface of the system-to-be should maintain the style of the existing system.
  (b3) The system-to-be should be secure.
- The domain expert in requirements engineering provides two constraints about security:
  (c1) To guarantee the security of the system-to-be, openness (or ease of extension) should not be considered.
  (c2) To improve the security of the system-to-be, the newest development techniques should be adopted.

The following predicates are used in [24] to formulate the requirements:

- the predicate Fash(int_f) is used to denote that the interface is fashionable;
- the predicate Open(sys) is used to denote that the system is open;
- the predicate New(sys) is used to denote that the system will be developed based on the newest techniques;
- the predicate Secu(sys) is used to denote that the system is secure.

Then the requirements above can be represented by a knowledge base

$$K_R = \left\{ \begin{array}{l} \mathrm{Fash(int\_f)},\ \mathrm{Open(sys)},\ \neg\mathrm{New(sys)},\ \neg\mathrm{Fash(int\_f)},\ \mathrm{Secu(sys)}, \\ \mathrm{Secu(sys)} \to \mathrm{New(sys)},\ \mathrm{Secu(sys)} \to \neg\mathrm{Open(sys)} \end{array} \right\}.$$

For simplicity of discussion, we abbreviate the knowledge base as

$$K_R = \{a1, a2, b1, b2, b3, c1, c2\}.$$

Evidently, K_R is inconsistent. The degree of responsibility of each requirement for the inconsistency is given as follows:

$$dr(K_R, a1) = dr(K_R, b2) = dr(K_R, b3) = \frac{1}{2},$$

$$dr(K_R, a2) = dr(K_R, b1) = dr(K_R, c1) = dr(K_R, c2) = \frac{1}{3}.$$

Allowing for the practical costs for abandoning requirements, developers are more interested in suggesting the stakeholders to change the requirements with the highest degree of responsibility. That is, developers are interested in either $\{a1, b3\}$ or $\{b2, b3\}$ at the beginning of negotiation with stakeholders.

Suppose that the user of the existing system refuses to change requirement $b3$. Instead, this stakeholder agrees to abandon requirement $b1$ after the first round of negotiation. Note that $dr(K_R, b1) = \frac{1}{3}$. This implies that stakeholders have to abandon two other requirements in order to restore the consistency of the set of requirements in the case that abandoning $b1$ is necessary. On the other hand, suppose that the domain expert insists to remain both the constraints $c1$ and $c2$ unchanged. In this case, developers are interested in recommending either $\{a1, b1, a2\}$ or $\{b2, b1, a2\}$ to stakeholders.

Suppose that the user of the existing system agrees to abandon requirement $b2$ as well after the second round of negotiation, while the seller agrees to withdraw requirement $a2$. Then requirements $\{b2, b1, a2\}$ are chosen as the ones to be abandoned. The revised set of requirements is represented by the following knowledge base

$$K_R^1 = \{a1, b3, c1, c2\}.$$

Now the set of requirements is consistent.

8. Conclusion

We have proposed an approach to identifying the degree of responsibility of each formula in a knowledge base for the inconsistency of that base in this paper. We proposed the measurement $dr(K, \alpha)$ for the degree of responsibility of a formula α in K for the inconsistency of K, which is exactly the minimal number of formulas in K that have to be removed from K in order to make the inconsistency of K counterfactually depend on α. Moreover, we have shown that the measurement $dr(K, \alpha)$ can be well explained in the context of causality and responsibility presented by Chockler and Halpern [1]. This means that the measurement $dr(K, \alpha)$ captures the degree of responsibility of α for the inconsistency indeed from the perspective of causality. Allowing for the duality of minimal inconsistent subsets and minimal correction subsets, we have shown that $dr(K, \alpha)$ can be expressed in terms of minimal correction subsets alternatively. Then we explored some intuitive logical properties satisfied by $dr(K, \alpha)$, and identified the complexity results for computing $dr(K, \alpha)$. Lastly, we use a small but explanatory example in requirements engineering to illustrate the application of the measurement dr.

On the other hand, some efficient algorithms for computing the minimal inconsistent subsets and minimal correction subsets have been proposed recently [33,38]. This provides a promising prospect for computing the degree of responsibility of each formula for the inconsistency in a given knowledge base efficiently.

Acknowledgements

The authors are grateful to anonymous reviewers for their valuable comments. This work was supported by the National Natural Science Foundation of China under Grant No. 61170300.

References

[1] H. Chockler, J.Y. Halpern, Responsibility and blame: a structural-model approach, J. Artif. Intell. Res. 22 (2004) 93–115.
[2] K. Mu, J. Hong, Z. Jin, W. Liu, From inconsistency handling to non-canonical requirements management: a logical perspective, Int. J. Approx. Reason. 54 (1) (2013) 109–131.
[3] K. Mu, Z. Jin, W. Liu, D. Zowghi, B. Wei, Measuring the significance of inconsistency in the viewpoints framework, Sci. Comput. Program. 78 (9) (2013) 1572–1599.
[4] D.P. Muiño, Measuring and repairing inconsistency in probabilistic knowledge bases, Int. J. Approx. Reason. 52 (6) (2011) 828–840.
[5] A. Hunter, S. Konieczny, Shapley inconsistency values, in: P. Doherty, J. Mylopoulos, C. Welty (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the 10th International Conference, KR06, AAAI Press, 2006, pp. 249–259.
[6] G. Qi, W. Liu, D. Bell, Measuring conflict and agreement between two prioritized belief bases, in: L.P. Kaelbling, A. Saffiotti (Eds.), Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence, IJCAI05, Edinburgh, UK, July 30–August 5, 2005, Professional Book Center, 2005, pp. 552–557.
[7] Y. Ma, G. Qi, P. Hitzler, Z. Lin, Measuring inconsistency for description logics based on paraconsistent semantics, in: Symbolic and Quantitative Approaches to Reasoning and Uncertainty, ECSQARU07, Hammamet, Tunisia, 31 October–November 2, 2007, in: Lect. Notes Comput. Sci., vol. 4724, Springer, 2007, pp. 30–41.
[8] K. McAreavey, W. Liu, P. Miller, K. Mu, Measuring inconsistency in a network intrusion detection rule set based on snort, Int. J. Semant. Comput. 5 (3) (2011) 281–322.
[9] K. Knight, Measuring inconsistency, J. Philos. Log. 31 (1) (2002) 77–98.
[10] K. Knight, Two information measures for inconsistent sets, J. Log. Lang. Inf. 12 (2) (2003) 227–248.
[11] D. Doder, M. Rašković, Z. Marković, Z. Ognjanović, Measures of inconsistency and defaults, Int. J. Approx. Reason. 51 (7) (2010) 832–845.
[12] J. Grant, Classifications for inconsistent theories, Notre Dame J. Form. Log. 19 (3) (1978) 435–444.
[13] L. Bertossi, A. Hunter, T. Schaub, Introduction to inconsistency tolerance, in: L. Bertossi, A. Hunter, T. Schaub (Eds.), Inconsistency Tolerance, in: Lect. Notes Comput. Sci., vol. 3300, Springer, 2004, pp. 1–16.
[14] J. Grant, A. Hunter, Measuring inconsistency in knowledge bases, J. Intell. Inf. Syst. 27 (2) (2006) 159–184.
[15] J. Grant, A. Hunter, Analysing inconsistent first-order knowledge bases, Artif. Intell. 172 (8–9) (2008) 1064–1093.
[16] A. Hunter, S. Konieczny, On the measure of conflicts: Shapley inconsistency values, Artif. Intell. 174 (14) (2010) 1007–1026.
[17] G. Xiao, Z. Lin, Y. Ma, G. Qi, Computing inconsistency measurements under multi-valued semantics by partial max-sat solvers, in: F. Lin, U. Sattler, M. Truszczynski (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the 10th International Conference, KR2010, AAAI Press, 2010, pp. 340–349.

[18] S. Konieczny, J. Lang, P. Marquis, Quantifying information and contradiction in propositional logic through epistemic actions, in: G. Gottlob, T. Walsh (Eds.), Proceedings of the 18th International Joint Conference on Artificial Intellignce, IJCAI2003, Morgan Kaufmann, 2003, pp. 106–111.

[19] A. Hunter, Logical comparison of inconsistent perspectives using scoring functions, Knowl. Inf. Syst. 6 (5) (2004) 528–543.

[20] A. Hunter, S. Konieczny, Measuring inconsistency through minimal inconsistent sets, in: G. Brewka, J. Lang (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Eleventh International Conference, KR08, AAAI Press, 2008, pp. 358–366.

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

[22] S. Jabbour, Y. Ma, B. Raddaoui, Inconsistency measurement thanks to mus decomposition, in: A.L.C. Bazzan, M.N. Huhns, A. Lomuscio, P. Scerri (Eds.), International conference on Autonomous Agents and Multi-Agent Systems, AAMAS'14, Paris, France, May 5–9, 2014, IFAAMAS/ACM, 2014, pp. 877–884.

[23] K. McAreavey, W. Liu, P.C. Miller, Computational approaches to finding and measuring inconsistency in arbitrary knowledge bases, Int. J. Approx. Reason. 55 (8) (2014) 1659–1693.

[24] K. Mu, W. Liu, Z. Jin, D. Bell, A syntax-based approach to measuring the degree of inconsistency for belief bases, Int. J. Approx. Reason. 52 (7) (2011) 978–999.

[25] S. Jabbour, B. Raddaoui, Measuring inconsistency through minimal proofs, in: L.C. van der Gaag (Ed.), Symbolic and Quantitative Approaches to Reasoning with Uncertainty – Proceedings of the 12th European Conference, ECSQARU 2013, Utrecht, The Netherlands, July 8–10, 2013, in: Lect. Notes Comput. Sci., vol. 7958, Springer, 2013, pp. 290–301.

[26] S. Jabbour, Y. Ma, B. Raddaoui, L. Sais, Prime implicates based inconsistency characterization, in: T. Schaub, G. Friedrich, B. O'Sullivan (Eds.), 21st European Conference on Artificial Intelligence, ECAI 2014, 18–22 August 2014, Prague, Czech Republic; Including Prestigious Applications of Intelligent Systems, PAIS 2014, in: Front. Artif. Intell. Appl., vol. 263, IOS Press, 2014, pp. 1037–1038.

[27] K. Mu, K. Wang, L. Wen, Approaches to measuring inconsistency for stratified knowledge bases, Int. J. Approx. Reason. 55 (2) (2014) 529–556.

[28] M. Thimm, Measuring inconsistency in probabilistic knowledge bases, in: Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence, UAI'09, AUAI Press, Montreal, Canada, June 2009, pp. 530–537.

[29] N. Potyka, Linear programs for measuring inconsistency in probabilistic logics, in: C. Baral, G.D. Giacomo, T. Eiter (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Fourteenth International Conference, KR 2014, Vienna, Austria, July 20–24, 2014, AAAI Press, 2014.

[30] M. Thimm, Inconsistency measures for probabilistic logics, Artif. Intell. 197 (2013) 1–24.

[31] A. Meliou, W. Gatterbauer, J.Y. Halpern, C. Koch, K.F. Moore, D. Suciu, Causality in databases, IEEE Data Eng. Bull. 33 (3) (2010) 59–67.

[32] J.Y. Halpern, J. Pearl, Causes and explanations: a structural-model approach. Part I: Causes, Br. J. Philos. Sci. 56 (4) (2005) 843–887.

[33] M.H. Liffiton, K.A. Sakallah, Algorithms for computing minimal unsatisfiable subsets of constraints, J. Autom. Reason. 40 (1) (2008) 1–33.

[34] R. Reiter, A theory of diagnosis from first principles, Artif. Intell. 32 (1) (1987) 57–95.

[35] G. Xiao, Y. Ma, Inconsistency measurement based on variables in minimal unsatisfiable subsets, in: L.D. Raedt, C. Bessière, D. Dubois, P. Doherty, P. Frasconi, F. Heintz, P.J.F. Lucas (Eds.), 20th European Conference on Artificial Intelligence, ECAI 2012, August 27–31, 2012, in: Front. Artif. Intell. Appl., vol. 242, IOS Press, Montpelier, France, 2012, pp. 864–869.

[36] H. Chockler, J.Y. Halpern, O. Kupferman, What causes a system to satisfy a specification? ACM Trans. Comput. Log. 9 (3) (2008) 20:1–20:26.

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

[38] J. Marques-Silva, F. Heras, M. Janota, A. Previti, A. Belov, On computing minimal correction subsets, in: F. Rossi (Ed.), Proceedings of the 23rd International Joint Conference on Artificial Intelligence, IJCAI 2013, Beijing, China, August 3–9, 2013, IJCAI/AAAI, 2013, pp. 615–622.