

Analysing Inconsistent Information using Distance-based Measures

John Grant* and Anthony Hunter†

April 25, 2016

Abstract

There have been a number of proposals for measuring inconsistency in a knowledgebase (i.e. a set of logical formulae). These include measures that consider the minimally inconsistent subsets of the knowledgebase, and measures that consider the paraconsistent models (3 or 4 valued models) of the knowledgebase. In this paper, we present a new approach that considers the amount each formula has to be weakened in order for the knowledgebase to be consistent. This approach is based on ideas of knowledge merging by Konienczny and Pino-Perez. We show that this approach gives us measures that are different from existing measures, that have desirable properties, and that can take the significance of inconsistencies into account. The latter is useful when we want to differentiate between inconsistencies that have minor significance from inconsistencies that have major significance. We also show how our measures are potentially useful in applications such as evaluating violations of integrity constraints in databases and for deciding how to act on inconsistency.

1 Introduction

Understanding the nature of inconsistency is an important topic if we are to develop autonomous systems that are able to behave intelligently with conflicting information. Although the early work of Grant in [Gra78] showed more than 30 years ago that it is possible to compare inconsistent sets of formulae, the great amount of research on measuring inconsistency occurred in the past decade. It turns out that there are different reasonable ways of measuring the inconsistency of a knowledgebase; these measures tend to be incompatible with one another in the sense that one measure assigns a larger inconsistency value to knowledgebase Δ than to Δ' while another does not.

The purpose of this paper is to introduce several inconsistency measures based on model distance. We work in propositional logic and assume that a knowledgebase contains only consistent formulae (i.e. each individual formula is consistent though the set of formulae may be inconsistent). This is a reasonable assumption as portions of conflicting information are typically consistent. However, we note that every inconsistent formula (other than the special case \perp) requires a conjunction; such a formula can always be split into consistent fragments. Every consistent formula has at least one model. We think of each model as a point in Euclidean space. The models of a knowledgebase are exactly the intersection of the set of models for each formula. When the knowledgebase is inconsistent, this intersection is empty.

In our method we use distance measures to measure the distances between models (points in space). The idea of our method is to dilate the points representing the models to regions of space in a minimal way so that the intersection of these regions is no longer empty. Our various proposals count different aspects of these dilations to come up with measures of inconsistency. Furthermore, this approach lends itself to assigning weights to atoms thereby capturing better

*Department of Computer Science and UMIACS, University of Maryland, College Park, MD 20742, USA

†Department of Computer Science, University College London, London, WC1E 6BT, UK

the significance of inconsistencies and provides new insight into the nature of inconsistency. For applications, it offers a better account for distances in the significance of parts of the knowledge that may be inconsistent. We illustrate how the new measures are potentially valuable tools for applications by considering violations of integrity constraints in databases.

The plan of this paper is as follows: (Section 2) We present the basic definitions and terminology; (Section 3) We define distance measures; (Section 4) We present the dilation of formulae; (Section 5) We present our definitions for measuring inconsistency using distance measures; (Section 6) We present our definitions for measuring information in inconsistent information; (Section 7) We show how weighting and costing can be used to take into account the significance of the information; (Section 8) We apply our approach to measure violations of integrity constraints; (Section 9) We compare our distance-based measures with several existing measures; (Section 10) We conclude the paper.

This paper is an extended version of [GH13]. We augment that paper by providing a systematic analysis of distance-based measures in terms of general properties of inconsistency measures, by showing how distance-based measures can be used for measuring information, by providing coverage of significance in terms of cost functions and cost rankings, and by providing a systematic comparison with key proposals for measures of inconsistency.

2 Preliminaries

We assume a propositional language \mathcal{L} of formulae composed of a finite set of atoms $\mathcal{A} = \{a_1, \dots, a_n\}$ as well as the logical connectives \wedge, \vee, \neg , and the punctuation symbols (and). Instead of subscripts for a we will often use consecutive letters; for instance (a, b, c) instead of (a_1, a_2, a_3) . A literal is an atom or a negated atom. We use ϕ and ψ for arbitrary formulae and α and β for literals. The set of atoms used on the composition of a formula ϕ is given by the function $\text{Atoms}(\phi)$. For example, $\text{Atoms}(\neg a \wedge (b \vee (\neg c \wedge d))) = \{a, b, c, d\}$.

A **knowledgebase** Δ is a finite set of *consistent* formulae. We let \vdash denote the classical **consequence relation**. Logical equivalence is defined in the usual way: $\Delta \equiv \Delta'$ iff $\Delta \vdash \Delta'$ and $\Delta' \vdash \Delta$. We find it useful to define also a stronger notion of equivalence as follows: knowledgebase Δ is **bijection-equivalent** to knowledgebase Δ' , denoted $\Delta \equiv_b \Delta'$ iff there is a bijection $f : \Delta \rightarrow \Delta'$ such that for all $\phi \in \Delta$, ϕ is logically equivalent to $f(\phi)$. For example, $\{a, b\}$ is logically equivalent but not bijection-equivalent to $\{a \wedge b\}$. We write $\mathbb{R}^{\geq 0}$ (resp. \mathbb{R}^+) for the set of nonnegative (resp. positive) real numbers and \mathcal{K} for the set of all knowledgebases.

Given a language \mathcal{L} we can assume an arbitrary sequence for the atoms \mathcal{A} , say (a_1, \dots, a_n) . Using this sequence of atoms, a **model** (i.e. an interpretation) is a sequence of 0s and 1s, written (b_1, \dots, b_n) where 0 means false and 1 means true for the corresponding atom. For $\phi \in \mathcal{L}$, $\text{Models}(\phi)$ denotes the set of interpretations for which ϕ is true using the usual evaluation of formulae in classical logic. For a knowledgebase Δ , $\text{Models}(\Delta)$ denotes the set of interpretations for which every $\phi \in \Delta$ is true. So if $\Delta = \{\phi_1, \dots, \phi_n\}$, then $\text{Models}(\Delta) = \text{Models}(\phi_1) \cap \dots \cap \text{Models}(\phi_n)$. We use $\mathcal{M}_{\mathcal{L}}$ to denote the set of models for the language \mathcal{L} (i.e. $\mathcal{M}_{\mathcal{L}}$ contains the 2^n sequences of (b_0, \dots, b_n) of 0s and 1s).

Often, instead of writing an interpretation as a sequence of 0s and 1s, we will write it as the (unique) binary number $b_1 \dots b_n$ (with leading 0s kept for easy readability). For example, if $\mathcal{A} = \{a_1, a_2, a_3\}$ and we assume the sequence of atoms (a_1, a_2, a_3) , then $\mathcal{M}_{\mathcal{L}} = \{000, 001, 010, 011, 100, 101, 110, 111\}$. Each interpretation can also be represented by a formula. For example, 101 can be represented by the formula $a_1 \wedge \neg a_2 \wedge a_3$.

We introduce a couple of subsidiary functions to analyse interpretations. For an interpretation m , let $\text{Digit}_i(m)$ return the i th digit of m (e.g. $\text{Digit}_2(1010) = 0$), and let $\text{Atom}_i(m)$ return the atom corresponding to the i th digit (e.g. if we assume the sequence of atoms (a, b, c, d) , then $\text{Atom}_2(1010) = b$).

We define the set of minimal inconsistent subsets of Δ , denoted $\text{MI}(\Delta)$, as follows (where for a set of formulae Γ , $\Gamma \vdash \perp$ denotes that Γ is inconsistent, and $\Gamma \not\vdash \perp$ denotes that Γ is consistent).

$$\text{MI}(\Delta) = \{\Sigma \mid \Sigma \subseteq \Delta \text{ and } \Sigma \vdash \perp \text{ and for all } \Sigma' \subset \Sigma, \Sigma' \not\vdash \perp\}$$

Any formula not involved in an inconsistency of a knowledgebase (i.e. it is not in a minimal inconsistent subset of the knowledgebase) is called a **free formula**. Thus the set of free formulae of a knowledgebase Δ is defined as follows.

$$\text{Free}(\Delta) = \{\alpha \in \Delta \mid \alpha \notin \bigcup \text{MI}(\Delta)\}$$

A more restricted notion than that of a free formula is the following notion: A formula α is a **safe formula** in a knowledgebase Δ when α has no atom in common with any other formula in Δ . Hence, the set of safe formulae is defined as follows. (Note that the usual definition of safe formula requires that α be consistent. We do not need this condition because our definition for a knowledgebase given above requires every formula in a knowledgebase to be consistent.)

$$\text{Safe}(\Delta) = \{\alpha \in \Delta \mid \text{Atoms}(\alpha) \cap \text{Atoms}(\Delta \setminus \{\alpha\}) = \emptyset\}$$

A safe formula cannot be involved in any inconsistency; hence every safe formula is free. However, a formula may be free but not safe: for example, a tautology that contains an atom that is also in another formula.

Next, we consider the concept of an inconsistency measure for knowledgebases. In the following definition of an inconsistency measure, the three constraints ensure that all and only consistent knowledgebases get measure 0, the measure is monotonic for subsets, and the removal of a safe formula leaves the measure unchanged.

Definition 1. *An inconsistency measure I assigns a nonnegative real value to every knowledgebase Δ . We assume three requirements for inconsistency measures as proposed in [HK10].*

- **Consistency** $I(\Delta) = 0$ iff Δ is consistent.
- **Monotony** If $\Delta \subseteq \Delta'$, then $I(\Delta) \leq I(\Delta')$.
- **Safe formula independence** If $\alpha \in \text{Safe}(\Delta)$, then $I(\Delta) = I(\Delta \setminus \{\alpha\})$.

In our definition we do not require free formula independence which is defined by substituting “Free” for “Safe” in Safe formula independence. However, we note that some authors (see for example [HK06]) require the stronger concept. We have chosen the weaker requirement in order not be constrained by the consideration of minimal inconsistent subsets as the basis for analysing conflict.

3 Distance measures

Given the set of models $\mathcal{M}_{\mathcal{L}}$ for a language \mathcal{L} , a distance measure, as defined next, is an assignment of a real number to each pair of models in $\mathcal{M}_{\mathcal{L}}$ that satisfies the usual axioms for a distance function (i.e. coincidence, symmetry, and subadditivity).

Definition 2. *For a given $\mathcal{M}_{\mathcal{L}}$, a **distance measure**, denoted d , is a function $d : \mathcal{M}_{\mathcal{L}} \times \mathcal{M}_{\mathcal{L}} \rightarrow \mathbb{R}^{\geq 0}$ satisfying the following three conditions:*

1. $d(m, m') = 0$ iff $m = m'$,
2. $d(m, m') = d(m', m)$,
3. $d(m, m') + d(m', m'') \geq d(m, m'')$.

For example, the function that assigns distance 1 between any two distinct models is a distance measure.

Definition 3. *A distance measure d is a **drastic measure** iff*

$$d(m, m') = \begin{cases} 1 & \text{if } m \neq m' \\ 0 & \text{if } m = m' \end{cases}$$

Next we define the Dalal measure using the concept of a contrary function.

Definition 4. *The contrary function, denoted $\text{Contrary} : \{0, 1\} \times \{0, 1\} \rightarrow \{0, 1\}$, is defined as follows*

$$\text{Contrary}(1, 1) = 0 \quad \text{Contrary}(1, 0) = 1 \quad \text{Contrary}(0, 1) = 1 \quad \text{Contrary}(0, 0) = 0$$

Definition 5. *Let \mathcal{L} be composed from n atoms, so that $\mathcal{M}_{\mathcal{L}}$ contains models with n digits. A distance measure d is a **Dalal measure** iff*

$$d(m, m') = \sum_{i=1}^n \text{Contrary}(\text{Digit}_i(m), \text{Digit}_i(m'))$$

A distance measure d is a Dalal measure [Dal88] when $d(m, m')$ is the number of digits that differ between m and m' . For a fixed n the Dalal measure is unique.

Example 1. *Consider the following measure for $n = 2$ which is a Dalal measure*

$$\begin{array}{cccc} d(11, 11) = 0 & d(11, 10) = 1 & d(11, 01) = 1 & d(11, 00) = 2 \\ d(10, 11) = 1 & d(10, 10) = 0 & d(10, 01) = 2 & d(10, 00) = 1 \\ d(01, 11) = 1 & d(01, 10) = 2 & d(01, 01) = 0 & d(01, 00) = 1 \\ d(00, 11) = 2 & d(00, 10) = 1 & d(00, 01) = 1 & d(00, 00) = 0 \end{array}$$

We now introduce the notion of a weighting function. We use a weighting function to assign a weight to each atom in a model. We write $w(i)$ for the weight of the i th atom. The idea is that the weight represents the significance of the atom and is used as a multiplicative factor. Since a distance cannot be negative, we do not deal with negative weights. Also, in order to assure the first property for a distance measure that we will define using a weighting function, we cannot allow 0 in the range of a weighting function.

Definition 6. *For \mathcal{L} composed from n atoms, a **weighting function** is a function $w : \{1, \dots, n\} \rightarrow \mathbb{R}^+$. We say that w is **uniform** iff for all $i \in \{1, \dots, n\}$, $w(i) = r$ for some $r \in \mathbb{R}^+$.*

Example 2. *Let $\mathcal{M}_{\mathcal{L}} = \{11, 10, 01, 00\}$. Then $w(1) = 0.5$ and $w(2) = 3$ is a weighting function.*

We define the concept of a weighted measure using a positive integer parameter k . As we will see, k is useful in distinguishing weighted measures. We will deal specifically only with the cases where $k = 1$ or $k = 2$ as they have a simple intuitive justification. Weighted measures could be defined more generally, but this will suffice for us. Basically, we want to make sure that the weight of each atom on which the models differ is applied and these weights are summed over all those atoms.

Definition 7. *For \mathcal{L} composed from n atoms, a **weighted measure** d is defined using a weighting function on the atoms and a positive integer parameter k :*

$$d(m, m') = \left(\sum_{i=1}^n [w(i) \times \text{Contrary}(\text{Digit}_i(m), \text{Digit}_i(m'))]^k \right)^{\frac{1}{k}}$$

Proposition 1. *Every weighted measure is a distance measure.*

Proof. The first two properties are obvious. The third property follows from the inequality:

$$\left(\sum_{i=1}^g A_i^k \right)^{\frac{1}{k}} + \left(\sum_{i=g+1}^n A_i^k \right)^{\frac{1}{k}} \geq \left(\sum_{i=1}^n A_i^k \right)^{\frac{1}{k}}$$

for $A_i \geq 0$ and $k \geq 1$. □

We start with the case where $k = 1$: the Manhattan measure.

Definition 8. Let \mathcal{L} be composed from n atoms. A weighted measure d is a **Manhattan measure** iff there is a weighting function w such that

$$d(m, m') = \sum_{i=1}^n w(i) \times \text{Contrary}(\text{Digit}_i(m), \text{Digit}_i(m'))$$

Example 3. Consider the following measure for $n = 2$ which is a Manhattan measure with the weighting function w where $w(1) = 3$ and $w(2) = 2$.

$$\begin{aligned} d(11, 11) &= 0 & d(11, 10) &= 2 & d(11, 01) &= 3 & d(11, 00) &= 5 \\ d(10, 11) &= 2 & d(10, 10) &= 0 & d(10, 01) &= 5 & d(10, 00) &= 3 \\ d(01, 11) &= 3 & d(01, 10) &= 5 & d(01, 01) &= 0 & d(01, 00) &= 2 \\ d(00, 11) &= 5 & d(00, 10) &= 3 & d(00, 01) &= 2 & d(00, 00) &= 0 \end{aligned}$$

In particular, a Dalal measure is a Manhattan measure with a uniform weighting function w where $w(i) = 1$ for each i . Another example of an important distance measure is the Euclidean measure, which treats space geometrically with $k = 2$.

Definition 9. Let \mathcal{L} be composed from n atoms. A distance measure d is a **Euclidean measure** iff there is a weighting function w such that

$$d(m, m') = \sqrt{\sum_{i=1}^n [w(i) \times \text{Contrary}(\text{Digit}_i(m), \text{Digit}_i(m'))]^2}$$

Example 4. Consider the following Euclidean measure where the weighting function is the same as in Example 3, that is, $w(1) = 3$ and $w(2) = 2$.

$$\begin{aligned} d(11, 11) &= 0.0 & d(11, 10) &= 2.0 & d(11, 01) &= 3.0 & d(11, 00) &= \sqrt{13} \\ d(10, 11) &= 2.0 & d(10, 10) &= 0.0 & d(10, 01) &= \sqrt{13} & d(10, 00) &= 3.0 \\ d(01, 11) &= 3.0 & d(01, 10) &= \sqrt{13} & d(01, 01) &= 0.0 & d(01, 00) &= 2.0 \\ d(00, 11) &= \sqrt{13} & d(00, 10) &= 3.0 & d(00, 01) &= 2.0 & d(00, 00) &= 0.0 \end{aligned}$$

Suppose we represent our n -digit models as points in n -dimensional space. Then the Manhattan measure (which involves following the edges of the hypercube) gives a greater distance between two points than the Euclidean measure (which takes the direct line between the two points). The Manhattan measure treats each side of the hypercube using its weight and adds the traversal of all of them. In contrast, the Euclidean measure discounts the distance with each further atom under consideration.

Example 5. Let $\mathcal{M}_{\mathcal{L}} = \{11, 10, 01, 00\}$ where d_D is the Dalal (Manhattan) measure while d_E is the corresponding Euclidean measure, and. $w(1) = w(2) = 1$. Consider the models 11 and 10. The Manhattan measure and the Euclidean measure are the same. Next, consider the models 11 and 00. The Euclidean measure in effect “discounts” the effect of the second digit being different between the models. We obtain

$$d_D(11, 11) = d_E(11, 11) < d_D(11, 10) = d_E(11, 10) < d_E(11, 00) < d_D(11, 00)$$

The Manhattan measure and the Euclidean measure are compatible with one another in the sense that the following hold.

$$\begin{aligned} d_D(m_1, m_2) &< d_D(m_3, m_4) \text{ iff } d_E(m_1, m_2) < d_E(m_3, m_4) \\ d_D(m_1, m_2) &= d_D(m_3, m_4) \text{ iff } d_E(m_1, m_2) = d_E(m_3, m_4) \end{aligned}$$

Furthermore, as we increase the parameter k , the distances become smaller. We included k to give a general definition for weighted measures.

Up to this point all the distance measures except for the drastic measure that we considered have been weighted measures. We end this section by presenting a distance measure that is not a weighted measure and is different from the drastic measure.

Example 6. *The following measure is not based on the atoms, and it is therefore not a weighted measure.*

$$\begin{array}{llll} d(11, 11) = 0 & d(11, 10) = 2 & d(11, 01) = 2 & d(11, 00) = 1 \\ d(10, 11) = 2 & d(10, 10) = 0 & d(10, 01) = 1 & d(10, 00) = 2 \\ d(01, 11) = 2 & d(01, 10) = 1 & d(01, 01) = 0 & d(01, 00) = 2 \\ d(00, 11) = 1 & d(00, 10) = 2 & d(00, 01) = 2 & d(00, 00) = 0 \end{array}$$

It is easy to check that the above example satisfies Definition 2. But in this case the distance between two models is smaller if they differ in both atoms than if they differ in one atom. Hence it is not a weighted measure.

4 Dilation of formulae

In order to define our new class of inconsistency measures we turn to the notion of dilation. Bloch and Lang, in [BL02], explore how some operations from mathematical morphology translate into a logical framework. One of the basic operations is the dilation of a set, which translates into the dilation of a formula (or its set of models). Essentially, for a formula ϕ , and a distance measure d , a dilation returns the models (or equivalently the formula specified by those models) that are at most a certain distance from (the models of) ϕ . The Dalal measure is a simple choice of distance measure to illustrate the idea. Suppose that ϕ is $a \wedge b$, and so the set of models is $\{11\}$. Using the Dalal measure, the set of dilations of distance 1 would be $\{11, 01, 10\}$, and so the resulting formula would be $a \vee b$. Then, the set of dilations of distance 2 would be $\{11, 01, 10, 00\}$, and so the resulting formula would be \top . Each dilation possibly weakens the previous formula in the sense that if ϕ is diluted to ϕ' then $\phi \vdash \phi'$. Note that dilation is inapplicable to any contradictory formula because such a formula has no models. That is why we do not allow contradictory formulae.

We harness this notion of dilation as follows.

Definition 10. *Let $\phi \in \mathcal{L}$ be a consistent propositional formula, let $k \in \mathbb{R}^{\geq 0}$, and let d be a distance measure. The set of **k-dilations** of ϕ with respect to d , $\text{Models}_d^k(\phi)$, is defined as follows:*

$$\text{Models}_d^k(\phi) = \{m \in \mathcal{M}_{\mathcal{L}} \mid \exists m' \in \text{Models}(\phi) \text{ such that } d(m', m) \leq k\}$$

We write $\text{Models}_D^k(\phi)$ for the special case where the distance measure d is the Dalal measure.

Hence, $\text{Models}_d^k(\phi)$ is the set of models whose distance (using d) is not more than k from some model of ϕ . The following result shows that the drastic measure is not sufficiently discriminating for our purposes since just a dilation of 1 will return all the models.

Proposition 2. *Let $\phi \in \mathcal{L}$ be a consistent propositional formula and let d be the drastic measure. For $k \geq 1$, $\text{Models}_d^k(\phi) = \mathcal{M}_{\mathcal{L}}$.*

Proof. For the drastic measure the distance of every model from the models of ϕ is 1. Hence, $\text{Models}_d^1(\phi) = \mathcal{M}_{\mathcal{L}}$. \square

Next, we extend Definition 10 to apply to sets of formulae. For this purpose it will be convenient to assume an arbitrary ordering, called the **standard ordering**, over the formulae in \mathcal{L} . The ordering has no significance. It just provides a standard way to put formulae into a sequence. Any ordering will do. For any $\Delta \subseteq \mathcal{L}$, we can then represent Δ as a tuple (ϕ_1, \dots, ϕ_n) , which we call the **standard form** of Δ , where $\Delta = \{\phi_1, \dots, \phi_n\}$ and $<$ is the standard ordering, and for each i , $1 \leq i < n$, $\phi_i < \phi_{i+1}$.

Definition 11. *Let (ϕ_1, \dots, ϕ_n) be the standard form of Δ , and let d be a distance measure. The set of **dilation profiles** with respect to d is*

$$\text{Profiles}_d(\Delta) = \{(k_1, \dots, k_n) \mid \text{Models}_d^{k_1}(\phi_1) \cap \dots \cap \text{Models}_d^{k_n}(\phi_n) \neq \emptyset\}$$

We write $\text{Profiles}_D(\Delta)$ when d is the Dalal measure.

The idea is to start with a sequence (ϕ_1, \dots, ϕ_n) of formulae in standard form, or equivalently, the sequence of their sets of models. Then, $\text{Profiles}_d(\Delta)$ is a sequence of numbers (k_1, \dots, k_n) such that the k_i -dilations of all the ϕ_i for $1 \leq i \leq n$ have a nonempty intersection. If we think of each k_i -dilation as the formula represented by the models, say ψ_i , then the nonempty intersection means that $\{\psi_1, \dots, \psi_n\}$ is consistent. We will be interested in minimizing $\text{Profiles}_d(\Delta)$ and using it to measure inconsistency.

Example 7. For $\Delta = \{a \wedge b, \neg a \wedge b\}$, and using the Dalal measure D ,

$$\text{Profiles}_D(\Delta) = \{(x, y) \mid x + y \geq 1\},$$

while using the non-weighted measure d' of Example 6,

$$\text{Profiles}_{d'}(\Delta) = \{(x, y) \mid x \geq 2 \text{ or } y \geq 2\}$$

Proposition 3. Let (ϕ_1, \dots, ϕ_n) be the standard form of a knowledgebase Δ and d a weighted measure with weighting function w .

1. $(0, \dots, 0) \in \text{Profiles}_d(\Delta)$ iff Δ is consistent.
2. If Δ is bijection-equivalent to Δ' and $(k_1, \dots, k_n) \in \text{Profiles}_d(\Delta)$ then there is a bijection $f : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ such that $(k_{f(1)}, \dots, k_{f(n)}) \in \text{Profiles}_d(\Delta')$.

Proof. 1. If Δ is consistent then Δ has a model, hence $(0, \dots, 0) \in \text{Profiles}_d(\Delta)$. Next, suppose that $(0, \dots, 0) \in \text{Profiles}_d(\Delta)$. So $k_1 = 0$ and ... and $k_n = 0$. This means that $\text{Models}_d^{k_1}(\phi_1) \cap \dots \cap \text{Models}_d^{k_n}(\phi_n) = \text{Models}(\Delta) \neq \emptyset$, meaning that Δ is consistent.

2. Let $\Delta = (\phi_1, \dots, \phi_n)$ and $\Delta' = (\phi'_1, \dots, \phi'_n)$ using the standard ordering. We cannot assume that $\phi_i \equiv \phi'_i$ in general; however, by bijection-equivalence, there must be a bijection $f : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ such that for each i , $1 \leq i \leq n$, $\phi'_i \equiv \phi_{f(i)}$. Suppose that $(k_1, \dots, k_n) \in \text{Profiles}_d(\Delta)$. Then $(k_{f(1)}, \dots, k_{f(n)}) \in \text{Profiles}_d(\Delta')$. □

In the next section, we will see examples using dilation. It will be useful there to deal only with minimal dilations that we next define.

Definition 12. A dilation profile $(k_1, \dots, k_n) \in \text{Profiles}_d(\Delta)$ is called **minimal** if and only if there is no dilation profile $(k'_1, \dots, k'_n) \in \text{Profiles}_d(\Delta)$ such that $(k_1, \dots, k_n) \neq (k'_1, \dots, k'_n)$ and $k'_i \leq k_i$ for all i , $1 \leq i \leq n$. We write $\text{Profiles}_d^{\text{min}}(\Delta)$ for the set of minimal dilation profiles.

So in Example 7 using the Dalal measure, $\text{Profiles}_D^{\text{min}}(\Delta) = \{(0, 1), (1, 0)\}$, and using the non-weighted measure $\text{Profiles}_{d'}^{\text{min}}(\Delta) = \{(0, 2), (2, 0)\}$.

5 Using dilation to measure inconsistency

Now we can use the set of minimal dilation profiles of a knowledgebase to assign it a measure of inconsistency. We define three measures. The first one sums a minimal sequence; the second picks the maximum value of a minimal sequence; while the third counts the number of nonzero values in a minimal sequence.

Definition 13. Let Δ be a knowledgebase and d a distance measure. The **d-sum inconsistency measure** is:

$$I_d^{\text{sum}}(\Delta) = \text{Min}\{x \mid \text{there exists } (k_1, \dots, k_n) \in \text{Profiles}_d^{\text{min}}(\Delta) \text{ such that } k_1 + \dots + k_n = x\}$$

We write I_D^{sum} if d is the Dalal measure.

Definition 14. Let Δ be a knowledgebase and d a distance measure. The **d-max inconsistency measure** is:

$$I_d^{max}(\Delta) = \text{Min}\{x \mid \text{there exists } (k_1, \dots, k_n) \in \text{Profiles}_d^{min}(\Delta) \text{ such that } \text{Max}\{k_1, \dots, k_n\} = x\}$$

We write I_D^{max} if d is the Dalal measure.

It is clear from the definitions that for all Δ , $I_d^{max}(\Delta) \leq I_d^{sum}(\Delta)$.

The third measure is somewhat different from the first two as it takes into account only the number of formulae that need to be diluted (hit) in order to make the set consistent. Intuitively, the more hits, the more inconsistent is the set of formulae. Hence, unlike the previous two cases, the magnitude of the dilation is not taken into account.

Definition 15. Let Δ be a knowledgebase and d a distance measure. The **d-hit inconsistency measure** is:

$$I_d^{hit}(\Delta) = \text{Min}\{x \mid \text{there exists } (k_1, \dots, k_n) \in \text{Profiles}_d^{min}(\Delta) \text{ such that } \text{Hit}(k_1, \dots, k_n) = x\}$$

where $\text{Hit}(k_1, \dots, k_n) = \sum_{i=1}^n z(k_i)$ and

$$z(k_i) = \begin{cases} 1 & \text{if } k_i > 0 \\ 0 & \text{if } k_i = 0 \end{cases}$$

We write I_D^{hit} in case d is the Dalal measure.

Before showing that these three definitions really define inconsistency measures, we give four examples. In these examples we use the Dalal measure.

Example 8. Let $\Delta_1 = \{a \wedge b, \neg a \wedge \neg b\}$. $\text{Profiles}_D^{min}(\Delta_1) = \{(1, 1), (2, 0), (0, 2)\}$. Hence, $I_D^{sum}(\Delta_1) = 2$, $I_D^{max}(\Delta_1) = 1$, and $I_D^{hit}(\Delta_1) = 1$.

k	$a \wedge b$	$\neg a \wedge \neg b$
0	{ 11 }	{ 00 }
1	{ 11, 10, 01 }	{ 10, 01, 00 }
2	{ 11, 10, 01, 00 }	{ 11, 10, 01, 00 }

Example 9. Let $\Delta_2 = \{a, \neg a \vee \neg b, b\}$. $\text{Profiles}_D^{min}(\Delta_2) = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. Hence, $I_D^{sum}(\Delta_2) = 1$, $I_D^{max}(\Delta_2) = 1$, and $I_D^{hit}(\Delta_2) = 1$.

k	a	$\neg a \vee \neg b$	b
0	{ 11, 10 }	{ 01, 10, 00 }	{ 11, 01 }
1	{ 11, 10, 01, 00 }	{ 11, 10, 01, 00 }	{ 11, 10, 01, 00 }

Example 10. Let $\Delta_3 = \{a \wedge b \wedge c, \neg a \wedge \neg b \wedge \neg c\}$. $\text{Profiles}_D^{min}(\Delta_3) = \{(1, 2), (2, 1), (3, 0), (0, 3)\}$. Hence, $I_D^{sum}(\Delta_3) = 3$, $I_D^{max}(\Delta_3) = 2$, and $I_D^{hit}(\Delta_3) = 1$.

k	$a \wedge b \wedge c$	$\neg a \wedge \neg b \wedge \neg c$
0	{ 111 }	{ 000 }
1	{ 111, 110, 101, 011 }	{ 010, 001, 100, 000 }
2	{ 111, 110, 101, 011, 100, 010, 001 }	{ 110, 101, 011, 010, 001, 100, 000 }
3	{ 111, 110, 101, 011, 100, 010, 001, 000 }	{ 111, 110, 101, 011, 010, 001, 100, 000 }

Example 11. Let $\Delta_4 = \{a, b, c, \neg a, \neg b, \neg c\}$. $\text{Profiles}_D^{min}(\Delta)$ includes $(1, 1, 1, 0, 0, 0)$, $(1, 1, 0, 0, 0, 1)$, $(1, 0, 0, 0, 1, 1)$, etc. Hence, $I_D^{sum}(\Delta) = 3$, $I_D^{max}(\Delta) = 1$, and $I_D^{hit}(\Delta) = 3$. We omit the table here because the second of the two rows is too long to include.

The following example illustrates the Euclidean measure. In this example we use integrity constraints to limit the set of models. There is no need to rewrite definitions: the integrity constraints just remove some models as we will explain.

Example 12. Consider knowledge-level sensor fusion with sensor1 and sensor2 observing the same object. Suppose sensor1 reports that “height is 1cm” and sensor2 reports that “height is 5cm”. Then the measure of inconsistency of this information is 1 unit to reflect that there is a disagreement in measurement. If we let a denote “height is 1cm”, and b denote “height is 5cm”, then $\Delta = \{a, b\}$, and $\text{Profiles}_E^{\min}(\Delta) = \{(1, 0), (0, 1)\}$ (assuming a and b are inconsistent together). Hence, $I_E^{\text{sum}}(\Delta) = 1$, $I_E^{\text{max}}(\Delta) = 1$, and $I_E^{\text{hit}}(\Delta) = 1$.

Now suppose sensor1 reports that “height is 1cm and width is 1cm” and sensor2 reports that “height is 5cm and width is 4cm”. We let a denote “height is 1cm”, b denote “height is 5cm”, c denote “width is 1cm”, and d denote “width is 4cm”. So sensor1 gives $a \wedge c$ and sensor2 gives $b \wedge d$. Thus, assuming that a is inconsistent with b , and c is inconsistent with d , we do not consider any of $\{1111, 1110, 1101, 1100, 1011, 0111, 0011\}$ as they denote inconsistent options. Therefore, $\text{Models}(a \wedge c) = \{1010\}$ and $\text{Models}(b \wedge d) = \{0101\}$. Hence, $\text{Models}_E^{\sqrt{2}}(a \wedge c) = \{1010, 0010, 0000, 0110, 1000, 1001\}$ and $\text{Models}_E^{\sqrt{2}}(b \wedge d) = \{0101, 0100, 0000, 0110, 0001, 1001\}$. Therefore, $\text{Models}_E^{\sqrt{2}}(a \wedge c) \cap \text{Models}_E^{\sqrt{2}}(b \wedge d) \neq \emptyset$. Hence, we obtain

$$\text{Profiles}_E^{\min}(\Delta) = \{(\sqrt{2}, \sqrt{2}), (1, \sqrt{3}), (\sqrt{3}, 1), (0, 2), (2, 0)\}$$

Therefore, $I_E^{\text{sum}}(\Delta) = 2$, $I_E^{\text{max}}(\Delta) = \sqrt{2}$, and $I_E^{\text{hit}}(\Delta) = 1$.

As promised earlier we show that the three inconsistency measures defined above satisfy the consistency, monotony, and free formula independence properties.

Proposition 4. *The d-sum inconsistency measure, the d-max inconsistency measure, and the d-hit inconsistency measure, each satisfy conditions 1 to 3 of Definition 1, and therefore all three are inconsistency measures.*

Proof. To show consistency it suffices to observe that $(0, \dots, 0) \in \text{Profiles}_d(\Delta)$ iff Δ is consistent. To show monotony let $\Delta \subseteq \Delta'$. It suffices to work with the case $\Delta' = \Delta \cup \{\phi\}$, where a single formula ϕ is added to Δ such that in the standard form ϕ comes after ϕ_n , and to consider only minimal dilations. Let $(k_1, \dots, k_n) \in \text{Profiles}_d^{\min}(\Delta)$. Then, for any $(k'_1, \dots, k'_n, k'_{n+1}) \in \text{Profiles}_d^{\min}(\Delta')$ $k'_i \geq k_i$ for $1 \leq i \leq n$ and $k'_{n+1} \geq 0$. Hence for each I_d^x measure, $I_d^x(\Delta) \leq I_d^x(\Delta')$. Finally, for safe formula independence, suppose that $\alpha \in \text{Safe}(\Delta)$. It suffices to deal only with the case where $\alpha = \phi_n$. Then $(k_1, k_2, \dots, k_{n-1}, 0) \in \text{Profiles}_d^{\min}(\Delta)$ because α need not be diluted. Hence, $(k_1, \dots, k_{n-1}) \in \text{Profiles}_d^{\min}(\Delta \setminus \{\alpha\})$ and for each I_d^x measure, $I_d^x(\Delta) = I_d^x(\Delta \setminus \{\alpha\})$. \square

Next we consider several properties that have been studied for inconsistency measures (see [HK06, HK10, MLJB11]) and show which of them hold for distance-based measures.

Definition 16. *The following are some of the properties considered for inconsistency measures:*

- **Free formula independence** *If $\alpha \in \text{Free}(\Delta)$, then $I(\Delta) = I(\Delta \setminus \{\alpha\})$.*
- **Dominance** *If $\{\alpha\} \vdash \beta$, and α is consistent, then $I(\Delta \cup \{\alpha\}) \geq I(\Delta \cup \{\beta\})$.*
- **MinInc** *If Δ is a minimal inconsistent set of formulae then $I(\Delta) = 1$.*
- **Attenuation** *For any two minimal inconsistent sets of formulae Δ_1 and Δ_2 if $|\Delta_1| > |\Delta_2|$ then $I(\Delta_1) < I(\Delta_2)$.*
- **Equal Conflict** *For any two minimal inconsistent sets of formulae Δ_1 and Δ_2 if $|\Delta_1| = |\Delta_2|$ then $I(\Delta_1) = I(\Delta_2)$.*
- **Maximal Contradiction** *If α is a contradiction, then $I(\{\alpha\}) \geq I(\Delta)$ for any minimal inconsistent set of formulae Δ .*
- **MinInc Separability** *If $\text{MI}(\Delta_1 \cup \Delta_2) = \text{MI}(\Delta_1) \cup \text{MI}(\Delta_2)$, and $\text{MI}(\Delta_1) \cap \text{MI}(\Delta_2) = \emptyset$, then $I(\Delta_1) + I(\Delta_2) = I(\Delta_1 \cup \Delta_2)$.*

These properties are not necessarily compatible with one another. Free formula independence states that the removal of a formula that does not participate in an inconsistency leaves the measure unchanged. Dominance states that if a consistent formula α logically implies formula β then the addition of α to a set of formulae cannot have smaller inconsistency than the addition of β . According to MinInc, all minimal inconsistent sets have measure 1. Attenuation gives an inverse relationship between the size of minimal inconsistent sets and their inconsistency; while Equal Conflict gives the same inconsistency to minimal inconsistent sets of the same size. Maximal contradiction states that the inconsistency of a single contradiction is maximal among minimal inconsistent sets. Finally, MinInc Separability means that if a set of formulae can be divided into two subsets in such a way that every minimal inconsistent subset of the union is in exactly one of the sets, then the inconsistency of the union is the sum of the inconsistencies of the subsets.

The next theorem shows which of the distance-based inconsistency measures have these properties.

Theorem 1. *The distance-based inconsistency measures I_d^{sum} , I_d^{max} , and I_d^{hit} have the following properties among the ones listed in Definition 16:*

Property	I_d^{sum}	I_d^{max}	I_d^{hit}
Free formula independence	×	×	✓
Dominance	✓	✓	✓
MinInc	×	×	✓
Attenuation	×	×	×
Equal Conflict	×	×	✓
Maximal Contradiction	×	×	×
MinInc Separability	×	×	×

Proof. (Free formula independence) For I_d^{sum} , consider $\Delta = \{a \wedge c, b \wedge \neg c, \neg a \vee \neg b\}$. As the only minimal inconsistent set contains the first two formulae, $\alpha \in \text{Free}(\Delta)$. Using the Dalal distance, $I_D^{sum}(\Delta \setminus \{\neg a \vee \neg b\}) = 1$, but $I_D^{sum}(\Delta) = 2$. For I_d^{max} , consider $\Delta = \{a \wedge b \wedge c \wedge d, \neg d \wedge e \wedge f \wedge g, \neg a \vee \neg e\} \wedge (\neg b \vee \neg f) \wedge (\neg c \vee \neg g)$. Just as in the previous case, the only minimal inconsistent set contains the first two formulae. Therefore, $(\neg a \vee \neg e) \wedge (\neg b \vee \neg f) \wedge (\neg c \vee \neg g) \in \text{Free}(\Delta)$. Using the Dalal distance, $I_D^{max}(\Delta \setminus \{(\neg a \vee \neg e) \wedge (\neg b \vee \neg f) \wedge (\neg c \vee \neg g)\}) = 1$, but $I_D^{max}(\Delta) = 2$. Finally, for I_d^{hit} , suppose that $\alpha \in \text{Free}(\Delta)$. It suffices to deal only with the case where $\alpha = \phi_n$. Then $(k_1, k_2, \dots, k_{n-1}, 0) \in \text{Profiles}_d^{min}(\Delta)$ because α need not be dilated. Hence, $(k_1, \dots, k_{n-1}) \in \text{Profiles}_d^{min}(\Delta \setminus \{\alpha\})$ so for the I_d^{hit} measure, $I_d^{hit}(\Delta) = I_d^{hit}(\Delta \setminus \{\alpha\})$.

(Dominance) Let $\{\alpha\} \vdash \beta$ where α is consistent. This implies that if $m \in \text{Models}(\{\alpha\})$ then $m \in \text{Models}(\{\beta\})$. Without loss of generality we assume that $\alpha \notin \Delta$ and $\beta \notin \Delta$. Let $(k_1, \dots, k_n) \in \text{Profiles}_d^{min}(\Delta \cup \{\alpha\})$. Then, $(k_1, \dots, k_n) \in \text{Profiles}_d^{min}(\Delta \cup \{\beta\})$; hence there must be $(k'_1, \dots, k'_n) \in \text{Profiles}_d^{min}(\Delta \cup \{\beta\})$ such that $k'_i \leq k_i$ for all i , $1 \leq i \leq n$. From this the result follows.

(MinInc) Let Δ be a minimal inconsistent set, $\Delta = \{\phi_1, \dots, \phi_n\}$ where, since Δ contains no contradiction, $n > 1$. It is always possible to weaken ϕ_1 to a tautology by some k-dilation, so that $(k, 0, \dots, 0) \in \text{Profiles}_d^{min}(\Delta)$; hence $I_d^{hit}(\Delta) = 1$. However, as shown in Example 10, I_d^{sum} and I_d^{max} need not equal 1.

(Attenuation) Consider Examples 8 and 9. Here $|\Delta_2| > |\Delta_1|$ but $I_D^{max}(\Delta_1) = I_D^{max}(\Delta_2)$ and $I_D^{hit}(\Delta_1) = I_D^{hit}(\Delta_2)$. Next, let $\Delta_5 = \{a, \neg a\}$. Here, $|\Delta_2| > |\Delta_5|$, but $I_D^{sum}(\Delta_2) = I_D^{sum}(\Delta_5)$.

(Equal Conflict) Consider Examples 8 and 10. $|\Delta_1| = |\Delta_3|$ but $I_D^{sum}(\Delta_1) \neq I_D^{sum}(\Delta_3)$ and $I_D^{max}(\Delta_1) \neq I_D^{max}(\Delta_3)$. In dealing with the MinInc property above we showed that for any minimal inconsistent set Δ , $I_d^{hit}(\Delta) = 1$, so, in particular, any two minimal inconsistent sets of the same size have identical I_d^{hit} value.

(Maximal Contradiction) As we do not allow a contradictory formula in Δ , this rule is not applicable here.

(MinInc Separability) Let $\Delta = \{a, b, c, \neg a \vee \neg b, \neg b \vee \neg c\}$ where $\Delta_1 = \{a, b, \neg a \vee \neg b\}$ and $\Delta_2 = \{b, c, \neg b \vee \neg c\}$. Thus, $\text{MI}(\Delta_1) = \{\Delta_1\}$ and $\text{MI}(\Delta_2) = \{\Delta_2\}$. Also, $\text{MI}(\Delta_1 \cup \Delta_2) = \text{MI}(\Delta) = \{\Delta_1, \Delta_2\}$. So the conditions that $\text{MI}(\Delta_1 \cup \Delta_2) = \text{MI}(\Delta_1) \cup \text{MI}(\Delta_2)$ and $\text{MI}(\Delta_1) \cap \text{MI}(\Delta_2) = \emptyset$ are satisfied. Using the Dalal measure we obtain the following.

- $I_D^{sum}(\Delta) = I_D^{max}(\Delta) = I_D^{hit}(\Delta) = 1$ because $(0, 1, 0, 0, 0) \in \text{Profiles}_D(\Delta)$.
- $I_D^{sum}(\Delta_1) = I_D^{max}(\Delta_1) = I_D^{hit}(\Delta_1) = 1$ because $(0, 1, 0) \in \text{Profiles}_D(\Delta_1)$.
- $I_D^{sum}(\Delta_2) = I_D^{max}(\Delta_2) = I_D^{hit}(\Delta_2) = 1$ because $(1, 0, 0) \in \text{Profiles}_D(\Delta_2)$.

Hence $I(\Delta) = I(\Delta_1) + I(\Delta_2)$ does not hold for any of these measures. \square

In order to compare two inconsistency measures, we define I_x and I_y to be *order-compatible* if for all knowledgebases Δ_1 and Δ_2 , $I_x(\Delta_1) < I_x(\Delta_2)$ iff $I_y(\Delta_1) < I_y(\Delta_2)$ and *order-incompatible* otherwise.

Proposition 5. *The d-sum inconsistency measure, the d-max inconsistency measure, and the d-hit inconsistency measure are pairwise order-incompatible.*

Proof. The result is illustrated using Examples 8 - 10. $I_D^{sum}(\Delta_2) < I_D^{sum}(\Delta_1)$, but $I_D^{max}(\Delta_2) = I_D^{max}(\Delta_1)$ as well as $I_D^{hit}(\Delta_2) = I_D^{hit}(\Delta_1)$. Finally, $I_D^{max}(\Delta_1) < I_D^{max}(\Delta_3)$, but $I_D^{hit}(\Delta_1) = I_D^{hit}(\Delta_3)$. \square

The d-sum inconsistency measure and the d-max inconsistency measure have been influenced by the definition for model-based merging operators by Konieczny and Pino Perez [KP98], and their dilation-based reformalization [GH08].

Dilation has a geometric interpretation using Euclidean distance in n-dimensional space. Consider the case with n atoms and weighting function w . Assign the point $(b_1 \cdot w(1), \dots, b_n \cdot w(n))$ to the model $b_1 \dots b_n$. For example, let $n = 3$ and weight function $w(1) = 2$, $w(2) = 5$, $w(3) = 4$. Then the model 101 is mapped to the point (2,0,4) and the model 110 is mapped to the point (2,5,0) (all points are in 3-dimensional space). For the distance between points (the models) we are using the Manhattan distance of moving along the edges of a hypercube, whereas the Euclidean distance is the “straight line” distance between the points.

Looking at the models this way as points in n-dimensional space using Euclidean distance, the k-dilation of a model is the set of points that represent models in a hypersphere of radius k with center at that point. As the k-dilation of a formula is the k-dilations of its models, geometrically, the k-dilation of a formula becomes the set of points that represent models in a union of hyperspheres. For the Manhattan distance substitute “hypercube” for “hypersphere”.

It is possible for two such hyperspheres or hypercubes to have a nonempty intersection that does not contain any models. Suppose that in the given example (1, 4, 2) is a point in the intersection. Such a point does not represent a model for the given weights. However, if we were using fractional truth values, the point would represent a model, namely with fractional truth values .5, .8, and .5 respectively for the atoms. We do not pursue this matter further in this paper.

6 Using Dilation to Measure Information

In this section we show how dilation can be used to measure information. In [GH11] we gave the following definition for an information measure.

Definition 17. *An information measure $J : \mathcal{K} \rightarrow \mathbb{R}^{\geq 0}$ is a function such that the following three conditions hold:*

1. *If $\Delta = \emptyset$ then $J(\Delta) = 0$.*
2. *If $\Delta' \subseteq \Delta$ and Δ is consistent, then $J(\Delta') \leq J(\Delta)$.*
3. *If Δ is consistent and $\exists \phi \in \Delta$ such that ϕ is not a tautology, then $J(\Delta) > 0$.*

This definition says nothing about inconsistent sets of formulae so there is a great deal of freedom in dealing with them. Later in this section we consider additional properties for information measures that take inconsistency into consideration.

In order to apply dilation we need to deal with the models of the formulae. Recall that for measuring inconsistency we defined a dilation profile, $\text{Profiles}_d(\Delta)$ (with respect to distance measure d), that gives a sequence of dilation values to attain consistency, that is, such that the intersection of the models becomes nonempty. For measuring information basically we move in the opposite direction with the models.

Let $\Delta = \{a \vee \neg a, b \vee \neg b\}$, so Δ contains only tautologies. $\text{Models}(\Delta) = \mathcal{M}_{\mathcal{L}} = \{11, 10, 01, 00\}$. Next, let $\Delta' = \{a \vee b\}$. Now, $\text{Models}(\Delta') = \{11, 10, 01\}$. So with more information there are fewer models. Finally, if $\Delta'' = \{a \wedge b\}$, then $\text{Models}(\Delta'') = \{11\}$ and there are fewer models still. Therefore, we define a measure of information on the basis that as the information increases, the set of models should decrease. We put this intuition in terms of a single dilation number.

In Definition 11 we extended the concept of the k -dilations of a formula to the concept of dilation profiles for a knowledgebase. We found this useful in measuring inconsistency. However, when we try to measure information, an alternative extension is more useful. We start by defining the set of k -dilations of Δ with respect to d as follows. For this definition we require Δ to be consistent.

Definition 18. Let Δ be a consistent knowledgebase, $k \in \mathbb{R}^{\geq 0}$, and d a distance measure. The set of k -dilations of Δ with respect to d is defined as follows:

$$\text{Models}_d^k(\Delta) = \bigcap_{\phi \in \Delta} \text{Models}_d^k(\phi)$$

The requirement that Δ be consistent ensures that $\text{Models}_d^k(\Delta) \neq \emptyset$ for any k . Next we define the dilation number concept leading to a single value.

Definition 19. Let Δ be a consistent knowledgebase and d a distance measure. The set of dilation numbers of Δ with respect to d is

$$\text{DilationNumber}_d(\Delta) = \{k \mid \text{Models}_d^k(\Delta) = \mathcal{M}_{\mathcal{L}}\}$$

The **minimal** dilation number w.r.t. d is $\text{DilationNumber}_d^{\text{min}}(\Delta) = \text{Min}\{k \mid k \in \text{DilationNumber}_d(\Delta)\}$. Again, D is used for the Dalal measure.

A further difference with dilation profiles is that there may be many minimal dilation profiles for an inconsistent knowledgebase but there is only one minimal dilation number for a consistent knowledgebase. In the example above, $\text{DilationNumber}_D^{\text{min}}(\Delta) = 0$, $\text{DilationNumber}_D^{\text{min}}(\Delta') = 1$, and $\text{DilationNumber}_D^{\text{min}}(\Delta'') = 2$. Now we are ready to define a distance-based information measure.

Definition 20. Let Δ be a knowledgebase and d a distance measure. The **d-information measure** is

$$J_d(\Delta) = \text{Max}\{k \mid \Delta' \text{ is a maximal consistent subset of } \Delta \text{ and } \text{DilationNumber}_d^{\text{min}}(\Delta') = k\}$$

As before, we write D for the Dalal measure.

In particular, if Δ is consistent, then $J_d(\Delta) = \text{DilationNumber}_d^{\text{min}}(\Delta)$. It is easy to verify that J_d is an information measure for any distance function d . Next we compute the information measures, using the Dalal distance measure, for several sets of formulae we considered previously.

Example 13. Consider the sets given in Examples 8 - 11. Then $J_D(\Delta_1) = J_D(\Delta_2) = 2$ and $J_D(\Delta_3) = J_D(\Delta_4) = 3$. Adding a consistent set $\Delta_5 = \{a \vee b \vee c\}$ we obtain $J_D(\Delta_5) = 1$.

We have previously considered several information measures in [GH11]. We just focus on one of these here, the one we called J_P . In order to define J_P we need a few definitions. A set of literals X is an **implicant** for Δ iff for each $\phi \in \Delta$, $X \vdash \phi$. A **proxy** for Δ is a minimal implicant of a maximal consistent subset of Δ . We write $\text{Proxies}(\Delta)$ for the set of proxies of Δ . For example, $\text{Proxies}(\{a, \neg a, b \vee c\}) = \{\{a, b\}, \{\neg a, b\}, \{a, c\}, \{\neg a, c\}\}$. The information measure J_P is then defined as the size of the largest proxy, that is, $J_P(\Delta) = \text{Max}\{|X| \mid X \in \text{Proxies}(\Delta)\}$. Clearly, no concept of dilation is used in the definition of J_P . Yet we obtain the following unexpected result.

Theorem 2. For all knowledgebases Δ , $J_P(\Delta) = J_D(\Delta)$.

Proof. As both J_P and J_D are defined as some maximal value over maximal consistent subsets, it suffices to consider only the case where Δ is consistent. If Δ contains only tautologies then $J_P(\Delta) = J_D(\Delta) = 0$. Next, let Δ contain the n propositional letters a_1, \dots, a_n and suppose that $J_P(\Delta) = r$ ($1 \leq r \leq n$). Without loss of generality let $A_r = \{a_1, \dots, a_r\}$ be a proxy of maximal size. Then $A_r \vdash \Delta$, but the relationship does not hold for any proper subset of A_r . Considering dilations, this means that r dilations are needed in order to obtain a model where each a_i , $1 \leq i \leq r$, is false while no dilations are needed for any a_i , $r < i \leq n$. Hence $J_D(\Delta) = r$. \square

Thus there is a strong connection between proxies and dilation for the Dalal distance measure, but not necessarily for other distance measures, for instance where different propositional letters get different weights. In [GH11] we also considered six possible properties for information measures given below.

Definition 21. Properties that an information measure J may satisfy:

- **Monotonic** If $\Delta \subseteq \Delta'$ then $J(\Delta) \leq J(\Delta')$.
- **Clarity** For all $\phi \in \Delta$, $J(\Delta) \geq J(\Delta \cup \{\neg\phi\})$.
- **Equivalence** If Δ is consistent and $\Delta \equiv \Delta'$, then $J(\Delta) = J(\Delta')$.
- **Bijection-Equivalence** If Δ and Δ' are bijection-equivalent then $J(\Delta) = J(\Delta')$.
- **Closed** If Δ is consistent and $\Delta \vdash \phi$, then $J(\Delta) = J(\Delta \cup \{\phi\})$.
- **Cumulative** If $\Delta \cup \{\phi\}$ is consistent and $\Delta \not\vdash \phi$, then $J(\Delta) < J(\Delta \cup \{\phi\})$.

In [GH11] we proved that J_P satisfies all the six properties given above. Hence J_D satisfies them as well. In fact, this result holds for other distance measures as well.

Proposition 6. For every distance measure d , J_d has the properties monotonic, clarity, equivalence, bijection-equivalence, closed, and cumulative.

Proof. These properties depend only on the sets of models of the formulae, not on the distances between models. Hence the result carries over from J_D to every J_d . \square

So in general, distance-based measures of information are well-behaved. Furthermore, given the equivalence between the proxy measure of information J_P and the Dalal measure J_D , distance-based measures provide some insight into the nature of existing proposals.

7 Significance

There are two reasons for presenting the distance-based measures of inconsistency in this paper. The first is to extend our understanding of the nature of inconsistency and how it can be measured. The second is to develop techniques for taking the significance of inconsistency into account. In this section we motivate the need to take significance into account and we add the concept of a cost function to deal with significance.

7.1 Need for taking significance into account

A simple way of taking significance into account is to assume a weighting function, and use a distance measure that takes this weight into account such as the Manhattan measure or the Euclidean measure, which we illustrate in the following example.

Example 14. Consider the following atoms where we may regard an inconsistency involving a as more significant than an inconsistency involving b .

- $a = \text{“There is rain in my city”}$
- $b = \text{“There is rain in a city 100Km from my city”}$

Consider the set of 2-digit models for (a, b) (i.e. the first digit refers to a , the second digit to b). Let $w(1) = 1$ and $w(2) = 0.1$ be the weighting function, and let d be the Manhattan distance.

Δ	$\{a \wedge b, \neg a \wedge \neg b\}$	$\{a \wedge b, \neg a \wedge b\}$	$\{a \wedge b, a \wedge \neg b\}$	$\{\neg a \wedge b, \neg a \wedge \neg b\}$
$I_d^{sum}(\Delta)$	1.1	1	0.1	0.1
$I_d^{max}(\Delta)$	1	1	0.1	0.1
$I_d^{hit}(\Delta)$	1	1	1	1

Using weights allows us to reduce inconsistency by applying an inconsistency resolution function (see [GH11]) that has maximal impact. For example, if $\Delta = \{a, \neg a, b, \neg b\}$ and $w(1)=1$, $w(2) = 10$, then deleting b or $\neg b$ reduces the inconsistency far better than deleting a or $\neg a$.

Whilst Example 14 shows how we can have different degrees of inconsistency based on significance, it does not take the context of the inconsistency into account. To illustrate what we mean by this, consider the following example where the measure is not a weighted measure.

Example 15. Consider the following atoms

- $a = \text{“There is an earthquake in my home town”}$
- $b = \text{“There is an electricity power cut in my home town”}$

In this situation, some assumptions we may have about the significance of inconsistency is as follows.

- if we have an inconsistency about whether or not there is an earthquake, then we have a very significant inconsistency.
- if we have an inconsistency about whether or not the electricity fails, then we have a moderate inconsistency.
- however, if we know that there is an earthquake, and there is an inconsistency about the electricity failing, then the significance of the inconsistency is low.

Consider the set of 2-digit models for (a, b) (i.e. the first digit refers to a , the second digit to b). We can capture this significance using the following distance measure which we also represent in Figure 1.

$$\begin{aligned}
d(11, 11) &= 0 & d(11, 10) &= 1 & d(11, 01) &= 9 & d(11, 00) &= 9 \\
d(10, 11) &= 1 & d(10, 10) &= 0 & d(10, 01) &= 9 & d(10, 00) &= 9 \\
d(01, 11) &= 9 & d(01, 10) &= 9 & d(01, 01) &= 0 & d(01, 00) &= 2 \\
d(00, 11) &= 9 & d(00, 10) &= 9 & d(00, 01) &= 2 & d(00, 00) &= 0
\end{aligned}$$

We illustrate the use of this distance measure with the following examples of knowledgebases.

Δ	$\{a \wedge b, \neg a \wedge \neg b\}$	$\{a \wedge b, \neg a \wedge b\}$	$\{a \wedge b, a \wedge \neg b\}$	$\{\neg a \wedge b, \neg a \wedge \neg b\}$
$I_d^{sum}(\Delta)$	9	9	1	2
$I_d^{max}(\Delta)$	9	9	1	2
$I_d^{hit}(\Delta)$	1	1	1	1

The difference between a weighted measure and a non-weighted measure is that for a weighted measure the atoms are independent of one another. That is not the case for non-weighted measures. So in Example 15 we can think of the 4 models as being in 2 groups: the group $\{11, 10\}$ and the group $\{00, 01\}$. Models within a group are close to one another but models in different groups have a larger distance. In that example the first atom is more important than the second atom; but the the second atom does not have a unique weight: its weight depends on the truth value of the first atom. However, if the groups are $\{11, 00\}$ and $\{01, 10\}$ then they are based on the sameness of the truth values of the atoms. With more atoms more groups can be formed.

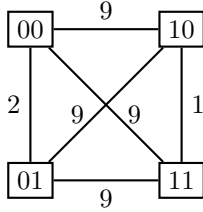


Figure 1: Consider Example 15. We can represent the distance between models using an undirected graph. Each node is a model, and the label on the arc is the distance between them.

7.2 Significance in terms of cost

We can measure the significance weight in monetary units. Given a set of inconsistencies, we can quantify how much we are prepared to pay to get the right answer. What we are prepared to pay depends on our context. But given a context, it is feasible to estimate what we are prepared to pay. To illustrate, suppose we are abroad on holiday. We may be prepared to pay quite a lot to resolve an inconsistency concerning the proposition “there is an earthquake in our home town”. We could cost it in terms of how much we are prepared to spend on expensive phone calls to resolve the inconsistency. Whereas, if the inconsistency is about something minor like “there is a power cut in our home town”, then we are much less likely to spend money calling home to resolve the contradiction.

We define a cost function based on the atoms \mathcal{A} in the language. We write $\rho(\mathcal{A})$ for the power set of \mathcal{A} .

Definition 22. A cost function is a function $C : \wp(\mathcal{A}) \rightarrow \mathbb{R}^{\geq 0}$ satisfying the following conditions:

1. $C(X) = 0$ iff $X = \emptyset$,
2. If $X \subseteq Y$ then $C(X) \leq C(Y)$,
3. $C(X) + C(Y) \geq C(X \cup Y)$.

The first condition states that the empty set is the only one to have 0 cost. The second condition states that the cost function is nondecreasing as the set membership increases. The third condition corresponds to the intuition that the cost of items taken together need not be as big as the sum of the individual costs.

The next example is a continuation of Example 15 but using a cost function. This cost function is closely related to the distance measure given there. We will investigate the relationship between distance measures and cost functions later.

Example 16. Let $a =$ “there is an earthquake in my home town” and $b =$ “there is an electricity power cut in my home town”. We could assign a cost function as follows: $C(\{a\}) = 9$, $C(\{b\}) = 1$, and $C(\{a, b\}) = 9$. So if the unit of cost is dollars, it could indicate the cost we would be prepared to pay to call home to resolve the inconsistency.

For inconsistencies in business databases, it is theoretically possible to quantify how much one would be prepared to resolve a set of inconsistencies (for example by paying for a lawyer to investigate). For example, if the inconsistency involves a bill owed by a client, then the cost one would be prepared to pay might be the value of the bill.

Example 17. Consider the following propositions.

- $a =$ “Invoice 1 states that company A owes our company \$2000”
- $b =$ “Invoice 2 states that company B owes our company \$3000”

- $c = \text{“Invoice 3 states that company } C \text{ owes our company } \4000”

We could assign a cost function as follows

$$\begin{array}{llll} C(\{a, b, c\}) = 9000 & C(\{a, b\}) = 5000 & C(\{a, c\}) = 6000 & C(\{b, c\}) = 7000 \\ C(\{a\}) = 2000 & C(\{b\}) = 3000 & C(\{c\}) = 4000 & \end{array}$$

So if the unit of cost is dollars, it could indicate the cost we would be prepared to pay lawyers to resolve the inconsistency.

Reducing significance to cost is similar to reducing utility to cost. The key to reducing it to a financial measure is that within the context of the application, the reduction is meaningful. For instance, above we consider the cost in terms of how much we would pay to resolve it. In other applications, it might be the cost of the potential damage if we don't resolve it.

7.3 Relating cost to distance

Now we consider how the cost function is related to the weighted and non-weighted distance measures. For this purpose we define a relation between models based on the atoms on which their truth value differs.

Definition 23. Let $X \subseteq \mathcal{A}$. The relation $\text{Diff}(X)$ relates models in the following manner:

$$\text{Diff}(X) = \{(m, m') \in \mathcal{M}_{\mathcal{L}} \times \mathcal{M}_{\mathcal{L}} \mid \text{for all } \alpha \in X, m \models \alpha \text{ iff } m' \not\models \alpha \\ \text{and for all } \beta \notin X, m \models \beta \text{ iff } m' \models \beta\}.$$

So $(m, m') \in \text{Diff}(X)$ means that m and m' disagree on the atoms in X and agree on all the atoms not in X . In particular, if $X = \emptyset$, then $\text{Diff}(X) = \{(m, m) \mid m \in \mathcal{M}_{\mathcal{L}}\}$, and for any $X \subseteq \mathcal{A}$, if $(m, m') \in \text{Diff}(X)$, then $(m', m) \in \text{Diff}(X)$. Furthermore, for any $X \subseteq \mathcal{A}$ and model m there is a unique m' such that $(m, m') \in \text{Diff}(X)$. In other words, each relation in $\text{Diff}(X)$ is in fact a function.

Example 18. Consider $\mathcal{A} = \{a\}$, and so $\mathcal{M}_{\mathcal{L}} = \{1, 0\}$.

X	$\text{Diff}(X)$
\emptyset	$\{(0, 0), (1, 1)\}$
$\{a\}$	$\{(1, 0), (0, 1)\}$

Example 19. Let $\mathcal{A} = \{a, b, c\}$. This means that $\mathcal{M}_{\mathcal{L}} = \{111, 110, 101, 100, 011, 010, 001, 000\}$.

1. If $X = \{a\}$ then $\text{Diff}(X)$ is the set containing the following tuples.

$$\begin{array}{cccc} (111, 011) & (110, 010) & (101, 001) & (100, 000) \\ (011, 111) & (010, 110) & (001, 101) & (000, 100) \end{array}$$

2. If $X = \{a, c\}$ then $\text{Diff}(X)$ is the set containing the following tuples.

$$\begin{array}{cccc} (111, 010) & (110, 011) & (101, 000) & (100, 001) \\ (010, 111) & (011, 110) & (000, 101) & (001, 100) \end{array}$$

We now show that the Diff relations form a partition of $\mathcal{M}_{\mathcal{L}} \times \mathcal{M}_{\mathcal{L}}$ in Propositions 7 and 8.

Proposition 7.

$$\bigcup_{X \subseteq \mathcal{A}} \text{Diff}(X) = \mathcal{M}_{\mathcal{L}} \times \mathcal{M}_{\mathcal{L}}$$

Proof. Consider any $(m, m') \in \mathcal{M}_{\mathcal{L}} \times \mathcal{M}_{\mathcal{L}}$. Let $X = \{\alpha \in \mathcal{A} \mid m \models \alpha \text{ iff } m' \not\models \alpha\}$. By definition, $(m, m') \in \text{Diff}(X)$. Therefore, for all $(m, m') \in \mathcal{M}_{\mathcal{L}} \times \mathcal{M}_{\mathcal{L}}$, there is an X such that $(m, m') \in \text{Diff}(X)$. Therefore, $\bigcup_{X \subseteq \mathcal{A}} \text{Diff}(X) = \mathcal{M}_{\mathcal{L}} \times \mathcal{M}_{\mathcal{L}}$. \square

Proposition 8. For all $X, Y \subseteq \mathcal{A}$, if $X \neq Y$, then $\text{Diff}(X) \cap \text{Diff}(Y) = \emptyset$.

Proof. Let $(m, m') \in \text{Diff}(X)$. Then, for all $\alpha \in X$, $m \models \alpha$ iff $m' \not\models \alpha$. If $X \subset Y$ then there is $\beta \in Y \setminus X$. So (m, m') cannot be in $\text{Diff}(Y)$ because $m \models \beta$ iff $m' \models \beta$. Otherwise, there is $\alpha \in X \setminus Y$. In this case (m, m') cannot be in $\text{Diff}(Y)$ because $m \models \alpha$ iff $m' \not\models \alpha$. \square

The following proposition shows that we can capture every weighted measure as a cost function. Thus cost functions subsume weighted measures as a representation.

Proposition 9. If d is a weighted measure, then there is a cost function C such that the following holds. For all $X \subseteq \mathcal{A}$, and for all $(m_1, m_2) \in \text{Diff}(X)$,

$$d(m_1, m_2) = C(X)$$

Proof. It suffices to show that for all X there is a value v such that if $(m_1, m_2) \in \text{Diff}(X)$ then $d(m_1, m_2) = v$, because in this case we can set $C(X) = v$. So suppose that $(m_1, m_2) \in \text{Diff}(X)$. This means that m_1 and m_2 disagree on all the literals in X and agree on the other literals. Let j_1, \dots, j_f be the positions of the literals in the first group, where m_1 and m_2 disagree. Thus, $d(m_1, m_2) = (\sum_{i=1}^f [w(j_i)]^k)^{\frac{1}{k}}$, which is the same for any such pair. We need to show that C is a cost function. This follows from the additive nature of weighted measure as we show now. The first property of a cost function follows from the fact that, as we showed earlier, $\text{Diff}(X) = \{(m, m) \mid \mathcal{M}_{\mathcal{L}} \times \mathcal{M}_{\mathcal{L}}\}$ iff $X = \emptyset$. To show the second property, let $X \subseteq Y$, $(m_1, m_2) \in \text{Diff}(X)$ and $(m_3, m_4) \in \text{Diff}(Y)$. Then $d(m_1, m_2) \leq d(m_3, m_4)$ because the set of atoms on which m_3 and m_4 differ includes the set of atoms on which m_1 and m_2 differ. For the third property consider that $C(X) = d(m_1, m_2)$ where m_1 and m_2 differ exactly on X , $C(Y) = d(m_2, m_3)$ where m_2 and m_3 differ exactly on Y , and $C(X \cup Y) = d(m_1, m_4)$ where m_1 and m_4 differ on $X \cup Y$. It suffices to consider the case where $X \cup Y$ is the largest it can be which occurs in case $X \cap Y = \emptyset$. But in this case $m_4 = m_3$ and the result follows from the third property of a distance measure and our earlier demonstration that a weighting measure is a distance measure. \square

The following example illustrates the proposition.

Example 20. Let $\mathcal{M} = \{111, 110, 101, 100, 011, 010, 001, 000\}$ for $\langle a, b, c \rangle$ (i.e. the first digit refers to a , the second to b , and the third to c). We can capture the Dalal measure d_D by using the cost function $C(X) = |X|$. We illustrate this as follows:

- Let $X = \{c\}$, then $\text{Diff}(X)$ is the set containing the following tuples.

$$\begin{array}{cccc} (111, 110) & (101, 100) & (011, 010) & (001, 000) \\ (110, 111) & (100, 101) & (010, 011) & (000, 001) \end{array}$$

Therefore, for all $(m_1, m_2) \in \text{Diff}(X)$, $d_D(m_1, m_2) = 1$ and $C(X) = 1$.

- Let $X = \{a, c\}$, then $\text{Diff}(X)$ is the set containing the following tuples.

$$\begin{array}{cccc} (111, 010) & (101, 000) & (011, 110) & (001, 100) \\ (010, 111) & (000, 101) & (110, 011) & (100, 001) \end{array}$$

Therefore, for all $(m_1, m_2) \in \text{Diff}(X)$, $d_D(m_1, m_2) = 2$ and $C(X) = 2$.

- Let $X = \{a, b, c\}$, then $\text{Diff}(X)$ is the set containing the following tuples.

$$\begin{array}{cccc} (111, 000) & (101, 010) & (011, 100) & (001, 110) \\ (000, 111) & (010, 101) & (100, 011) & (110, 001) \end{array}$$

Therefore, for all $(m_1, m_2) \in \text{Diff}(X)$, $d_D(m_1, m_2) = 3$ and $C(X) = 3$.

We now consider an example to show that the converse of Proposition 9 does not hold; that is, there is not necessarily a weighted measure that corresponds to a given cost function.

Example 21. Consider Example 16 where \mathcal{A} and C are defined. Suppose that there is a weighted measure d that corresponds to C . Then, for all $X \subseteq \mathcal{A}$ and $(m_1, m_2) \in \text{Diff}(X)$, $d(m_1, m_2) = C(X)$. Note that $(01, 11) \in \text{Diff}(\{a\})$, so $d(01, 11) = 9$. This forces $w(1) = 9$. Next, $(00, 01) \in \text{Diff}(\{b\})$, so $d(00, 01) = 1$. This forces $w(2) = 1$. But $(00, 11) \in \text{Diff}(\{a, b\})$, so $d(00, 11) = 9$. Since d is a weighted measure, $d(00, 11) = ([w(1)]^k + [w(2)]^k)^{\frac{1}{k}} = (9^k + 1^k)^{\frac{1}{k}}$ for some positive integer k , yielding the equation $(9^k + 1^k)^{\frac{1}{k}} = 9$ which is incorrect for all such k .

However, the following result shows that any cost function can always be captured by a distance measure (which need not be a weighted measure).

Proposition 10. If C is a cost function, then there is a distance measure d such that the following holds: For all $X \subseteq \mathcal{A}$, and for all $(m_1, m_2) \in \text{Diff}(X)$,

$$d(m_1, m_2) = C(X)$$

Proof. We need to show that the d defined above is a distance measure. The first property follows from the first property of a cost function. The second property follows from the symmetry of every $\text{Diff}(X)$ as the order of m and m' is irrelevant. For the third property, consider any 3 models m , m' , and m'' . Recall that $d(m, m') = C(X)$ where m and m' differ on X , $d(m', m'') = C(Y)$ where m' and m'' differ on Y , and $d(m, m'') = C(Z)$ where m and m'' differ on Z . By property 2 of a cost measure, $d(m, m'')$ is largest when $Z = X \cup Y$ and $X \cap Y = \emptyset$. The result follows from property 3 of a cost function. \square

We now consider an example to show that the converse of Proposition 10 does not hold; that is, there is not necessarily a cost function corresponding to a given distance measure.

Example 22. We recall Example 15 that we represented in Figure 1. Let $\mathcal{A} = \{a, b\}$. For the non-weighted distance measure we have $d(00, 01) = 2$ and $d(10, 11) = 1$. For each of these, the conflict between these models is b , so they are both in $\text{Diff}(\{b\})$. By the way that a cost function is defined, we obtain both $C(\{b\}) = 1$ and $C(\{b\}) = 2$ which shows that a cost function cannot be defined.

Weighted measure is the simplest option for introducing significance into distance-based inconsistency measures. However, there appears to be a need for a more sophisticated formalization such as offered by non-weighted measures. The downside of a non-weighted measure is the need to specify the distance for each pair of models. An alternative option is a cost function which offers an intermediate solution between weighted measure and non-weighted distance measure. It is intermediate in that it is more expressive than weighted measure but less expressive than non-weighted measure, and it is a more efficient representation than the non-weighted measure. We formalize these concepts using the following definition and proposition.

Definition 24. A **cost statement** is of the form $C(X) = v$ where $X \subseteq \mathcal{A}$ and $v \in \mathbb{R}^{\geq 0}$; a **weighted distance statement** is of the form $w(a) = v$ where $a \in \mathcal{A}$ and $v \in \mathbb{R}^+$; and a **non-weighted distance statement** is of the form $d(m, m') = v$ where $m, m' \in \mathcal{M}$ and $v \in \mathbb{R}^{\geq 0}$.

Proposition 11. The following give the relative costs of the three approaches.

1. A weighted measure d is specified by $|\mathcal{A}|$ weighted distance statements.
2. A cost function C is specified by $2^{|\mathcal{A}|} - 1$ cost statements.
3. A non-weighted distance measure d is specified by $2^{|\mathcal{A}|-1} \times (2^{|\mathcal{A}|} - 1)$ non-weighted distance statements.

Proof. (1) The $|\mathcal{A}|$ weights (along with the parameter k) completely specify a weighted measure. (2) The cost of every nonempty set of atoms must be given. (3) There are $2^{|\mathcal{A}|}$ interpretations. Using the facts that $d(m, m) = 0$ and $d(m, m') = d(m', m)$ means that the number of distances that must be given is the binomial coefficient $\binom{2^{|\mathcal{A}|}}{2}$, which is $2^{|\mathcal{A}|-1} \times (2^{|\mathcal{A}|} - 1)$. \square

As shown in this proposition, an exponential number of statements are needed to specify a cost function. We now give a condition for cost functions that reduces the complexity to linear size.

Definition 25. A cost function C is **additive** iff for all $X, Y \subseteq \mathcal{A}$, if $X \cap Y = \emptyset$, then $C(X \cup Y) = C(X) + C(Y)$.

Proposition 12. An additive cost function C is specified by $|\mathcal{A}|$ cost statements.

Proof. For an additive cost function it suffices to specify the cost of singleton subsets of \mathcal{A} . \square

Example 23. Consider the following cost function which satisfies the additive property.

$$C(\{a\}) = 1 \quad C(\{b\}) = 2 \quad C(\{a, b\}) = 3$$

Hence, we get the following distances between models.

$$\begin{aligned} d(11, 11) &= 0 & d(11, 10) &= 2 & d(11, 01) &= 1 & d(11, 00) &= 3 \\ d(10, 11) &= 2 & d(10, 10) &= 0 & d(10, 01) &= 3 & d(10, 00) &= 1 \\ d(01, 11) &= 1 & d(01, 10) &= 3 & d(01, 01) &= 0 & d(01, 00) &= 2 \\ d(00, 11) &= 3 & d(00, 10) &= 1 & d(00, 01) &= 2 & d(00, 00) &= 0 \end{aligned}$$

Using an additive cost function is the same as using a special type of weighted measure, namely a Manhattan measure, as we show in the following proposition. In other words, for each $\alpha \in \mathcal{A}$, we equate $C(\{\alpha\})$ with the weight assigned to α by the weighting function.

Proposition 13. If C is an additive cost function, and d is a distance measure such that for all $X \subseteq \mathcal{A}$, and for all $(m, m') \in \text{Diff}(X)$,

$$d(m, m') = C(X),$$

then d is a Manhattan measure.

Proof. We already showed in Proposition 10 that if C is a cost function and d is defined as above, then d is a distance measure. For the new result let C be an additive cost function. Writing $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$, $d(m, m') = C(X) = \sum_{\alpha_i \in X} C(\{\alpha_i\})$. Hence d is a Manhattan measure with $w(i) = C(\{\alpha_i\})$. \square

Cost functions offer a useful intermediate choice between the full expressibility of non-weighted measures and weighted measures. The choice of representation to use, trading expressibility for efficiency, depends on the application.

7.4 Cost ranking

In some applications, actual costs may be unknown but there is comparative information about costs. For instance, if X and Y are sets of atoms we may know that the cost of X is less than the cost of Y without knowing the exact costs. Now we investigate such cost ranking. We define a cost ranking as a binary relation \leq_R and in the usual way write $X <_R Y$ to denote $X \leq_R Y$ and $Y \not\leq_R X$; $X \simeq_R Y$ to denote $X \leq_R Y$ and $Y \leq_R X$; and let the binary relation $X \geq_R Y$ be equivalent to $Y \leq_R X$.

Definition 26. A **cost ranking** is a binary relation \leq_R over $\wp(\mathcal{A})$ that has the following properties: (for all X and Y)

- (Reflexive) $X \simeq_R X$
- (Transitive) $X \leq_R Y$ and $Y \leq_R Z$ implies $X \leq_R Z$
- (Monotonic) $X \subseteq Y$ implies $X \leq_R Y$

A **linear cost ranking** has the additional property: $X <_R Y$ or $X \simeq_R Y$ or $Y <_R X$.

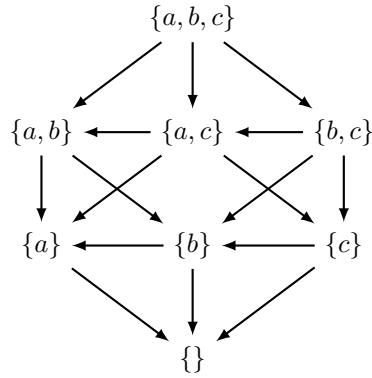
$X <_R Y$ means that we are prepared to pay more to resolve the inconsistencies involving all the atoms in Y than to resolve the inconsistencies involving all the atoms in X .

Next we revisit Example 17 but now we use cost ranking instead of a cost function.

Example 24. Consider the following propositions.

- $a =$ “Invoice 1 states that company A owes our company \$2000”
- $b =$ “Invoice 2 states that company B owes our company \$3000”
- $c =$ “Invoice 3 states that company C owes our company \$4000”

We could represent the binary relation as follows where for sets X and Y , the arrow from Y to X denotes $X \leq_R Y$.



We may go further and assume a linear ordering over the sets as follows.

$$\{\} \longleftarrow \{a\} \longleftarrow \{b\} \longleftarrow \{c\} \longleftarrow \{a, b\} \longleftarrow \{a, c\} \longleftarrow \{b, c\} \longleftarrow \{a, b, c\}$$

In terms of \leq_R this means that

$$\emptyset \leq_R \{a\} \leq_R \{b\} \leq_R \{c\} \leq_R \{a, b\} \leq_R \{a, c\} \leq_R \{b, c\} \leq_R \{a, b, c\}$$

In this case we obtained a linear cost ranking from a cost function. The following result shows that this can always be done.

Proposition 14. For each cost function C , there is a linear cost ranking \leq_C over $\wp(\mathcal{A})$ such that the following holds:

$$\text{For all } X, Y \subseteq \mathcal{A}, X \leq_C Y \text{ iff } C(X) \leq C(Y)$$

Proof. We define $X \leq_C Y$ iff $C(X) \leq C(Y)$. The result follows from the monotonicity of the cost function (Property (2)) and properties of real numbers. \square

As we will see, cost ranking can be finer grained than a weighted measure. Furthermore, cost ranking is more flexible in that it can be a partial ordering, whereas a distance measure imposes a linear ordering over the models of the atoms.

Next we consider how cost ranking is related to weighted and non-weighted distance measures. The following result shows that the inequality derived from a weighted measure can be captured by a linear cost ranking.

Proposition 15. For every weighted measure d there is a linear cost ranking \leq_R over $\wp(\mathcal{A})$ such that the following holds where $(m_1, m_2) \in \text{Diff}(X)$ and $(m_3, m_4) \in \text{Diff}(Y)$:

$$X \leq_R Y \text{ iff } d(m_1, m_2) \leq d(m_3, m_4)$$

Proof. First we observe that for every weighted measure d if $(m_1, m_2) \in \text{Diff}(X)$ and $(m'_1, m'_2) \in \text{Diff}(X)$ then $d(m_1, m_2) = d(m'_1, m'_2)$ because every weighted measure does the same computation to obtain $d(m_1, m_2)$ and $d(m'_1, m'_2)$. Thus \leq_R is well-defined and we need to show that it is a linear cost ranking. The argument for \leq_R being well-defined gives reflexivity immediately. Transitivity follows from the transitivity of \leq for real numbers and linearity follows from the complete ordering of the reals. Monotonicity follows from the way that a weighted measure is defined by the addition of weights for the atoms on which the models differ. \square

Example 25. Consider the weighted measure of Example 3. The corresponding cost ranking \leq_R over $\wp(\mathcal{A})$ is

$$\emptyset <_R \{b\} <_R \{a\} <_R \{a, b\}$$

The following example shows that going in the other direction is not always possible. That is, there need not be a weighted measure corresponding to a given linear cost ranking.

Example 26. Consider $\mathcal{A} = \{a, b\}$ where the linear cost ranking \leq_R is

$$\emptyset <_R \{a\} \simeq_R \{b\} \simeq_R \{a, b\}$$

For any possible weighted measure corresponding to \leq_R the following equations must be true:

$$([w(a)]^k)^{\frac{1}{k}} = ([w(b)]^k)^{\frac{1}{k}} = ([w(a) + w(b)]^k)^{\frac{1}{k}}$$

But this can work only for $w(a) = w(b) = 0$, which is not allowed.

Although in general there is no weighted measure that corresponds to a given linear cost ranking, we now show that there is such a distance measure.

Proposition 16. For every linear cost ranking \leq_R over $\wp(\mathcal{A})$ there is a distance measure d such that the following holds where $(m_1, m_2) \in \text{Diff}(X)$ and $(m_3, m_4) \in \text{Diff}(Y)$

$$d(m_1, m_2) \leq d(m_3, m_4) \text{ iff } X \leq_R Y$$

Proof. Use \leq_R to order all subsets of \mathcal{A} as

$$\emptyset = X_0 <_R X_1 \leq_R X_2 \leq_R \dots \leq_R X_m$$

Assign a real value r_i to every X_i in order, starting with X_0 in such a way that $r_0 = 0$, $r_i < r_{i+1}$ in case $X_i < X_{i+1}$, $r_i = r_{i+1}$ in case $X_i \simeq X_{i+1}$, and also $r_i \leq r_j + r_k$ in case $X_j \cup X_k = X_i$ and $X_j \cap X_k = \emptyset$ with $j < i$ and $k < i$. There is a problem with this process in case $r_{i-1} \geq r_j + r_k$, in which case first r_{i-1} must be reduced so that $r_{i-1} \leq r_j + r_k$. In fact, all $r_{i-\ell}$ for which $r_{i-\ell} \geq r_j + r_k$ must be reduced appropriately. The density of the real numbers allows for this finite process.

Next, define d such that $d(m, m') = r_i$ whenever $(m, m') \in \text{Diff}(X_i)$. We need to show that d is a distance function. The first property follows from the fact that $(m, m') \in \text{Diff}(X_0)$ iff $m' = m$. The second property follows from the symmetry of $\text{Diff}(X)$. We must still show that $d(m, m') + d(m', m'') \geq d(m, m'')$. So let $(m, m') \in \text{Diff}(X_j)$, $(m', m'') \in \text{Diff}(X_k)$, and $(m, m'') \in \text{Diff}(X_i)$. Then by the definition of d , $d(m, m') = r_j$, $d(m', m'') = r_k$, and $d(m, m'') = r_i$. Also, $X_i \subseteq X_j \cup X_k$. Therefore, by the construction, $r_i \leq r_j + r_k$, which proves the third property. \square

We illustrate Proposition 16 in the next examples.

Example 27. Returning to Example 26, we use the following ordering: $X_0 = \emptyset$, $X_1 = \{a\}$, $X_2 = \{b\}$, and $X_3 = \{a, b\}$. Also, let $r_0 = 0$, $r_1 = 1$, $r_2 = 1$, and $r_3 = 1$. We obtain the following table listing the model pairs for all $\text{Diff}(X)$. Therefore, for all (m, m') , if $m \neq m'$, then $d(m, m') = 1$,

X	$\text{Diff}(X)$
\emptyset	$\{(11, 11), (10, 10), (01, 01), (00, 00)\}$
$\{a\}$	$\{(11, 01), (10, 00), (01, 11), (00, 10)\}$
$\{b\}$	$\{(11, 10), (10, 11), (01, 00), (00, 01)\}$
$\{a, b\}$	$\{(11, 00), (10, 01), (01, 10), (00, 11)\}$

Example 28. For the linear cost ranking in Example 24 we define the following: $X_0 = \emptyset$, $X_1 = \{a\}$, $X_2 = \{b\}$, $X_3 = \{c\}$, $X_4 = \{a, b\}$, $X_5 = \{a, c\}$, $X_6 = \{b, c\}$, and $X_7 = \{a, b, c\}$. Then, let $r_0 = 0$, $r_1 = 1.0$, $r_2 = 1.1$, $r_3 = 1.2$, $r_4 = 1.3$, $r_5 = 1.4$, $r_6 = 1.5$, and $r_7 = 1.6$. For instance, $d(101, 111) = 1.1$ because $(101, 111) \in \text{Diff}(\{b\})$ and $d(111, 010) = 1.4$ because $(111, 010) \in \text{Diff}(\{a, c\})$. Note how, for instance, $d(101, 111) + d(111, 010) = 2.5 > d(101, 010) = 1.6$.

But suppose that initially we made different choices. Let's start with $r_0 = 0$, $r_1 = 1$, $r_2 = 2$, and $r_3 = 3$. Now there is a problem when we try to assign a value to r_4 because we must have $r_4 > 3$ and also $r_4 \leq r_1 + r_2 = 3$. The only number we have to reduce is r_3 as $r_2 < 3$. So we can change r_3 from 3 to 2.5 for instance. Now we can assign $r_4 = 3$ and there is no problem. We can continue, for example, with $r_5 = 3.5$, $r_6 = 4$, and $r_7 = 5$.

Our final example in this subsection illustrates that going in the opposite direction in Proposition 16 is not always possible. That is, for a given distance measure there need not be a corresponding linear cost ranking.

Example 29. Consider the distance measure given below for an example with two atoms:

$$\begin{array}{cccc} d(11, 11) = 0 & d(11, 10) = 2 & d(11, 01) = 1 & d(11, 00) = 2 \\ d(10, 11) = 2 & d(10, 10) = 0 & d(10, 01) = 2 & d(10, 00) = 2 \\ d(01, 11) = 1 & d(01, 10) = 2 & d(01, 01) = 0 & d(01, 00) = 1 \\ d(00, 11) = 2 & d(00, 10) = 2 & d(00, 01) = 1 & d(00, 00) = 0 \end{array}$$

Note first that $d(01, 11) = 1 < d(10, 11) = 2$ forces the corresponding cost ranking \leq_R to have $\{a\} <_R \{b\}$. But then consider that $d(01, 00) = 1 < d(00, 10) = 2$ which forces $\{b\} <_R \{a\}$. Hence no cost ranking is possible.

In conclusion, cost rankings offer an alternative to cost functions in helping to define distance measures. They are richer in some respects, but they are also weaker in that they provide less explicit information.

8 Measuring violations of integrity constraints

In this section we consider measuring violations of integrity constraints in knowledgebases. As integrity constraints must be satisfied, we slightly revise our definitions so that only the data is diluted and not the integrity constraints. We assume that relational data, Δ , and integrity constraints, Γ are treated as propositional formulae.

Definition 27. Let (ϕ_1, \dots, ϕ_n) be the standard form of a set of consistent propositional formulae Δ . Let $\Gamma \subseteq \mathcal{L}$ be a consistent set of propositional formulae, and let d be a distance measure. The set of **dilation profiles** with respect to d and Γ is as follows:

$$\text{Profiles}_d(\Delta, \Gamma) = \{(k_1, \dots, k_n) \mid \text{Models}_d^{k_1}(\phi_1) \cap \dots \cap \text{Models}_d^{k_n}(\phi_n) \cap \text{Models}(\Gamma) \neq \emptyset\}$$

Minimal dilation profiles are defined analogously to Definition 12 with $\text{Profiles}_d(\Delta, \Gamma)$ substituted for $\text{Profiles}_d(\Delta)$. Similarly, $I_d^{\text{sum}}(\Delta, \Gamma)$, $I_d^{\text{max}}(\Delta, \Gamma)$, and $I_d^{\text{hit}}(\Delta, \Gamma)$, are defined analogously to Definitions 13–15.

In order to render the following examples more concise, we will represent the integrity constraints by a scheme. The variables of the scheme are instantiated by the constants appearing in the database. We will use upper case letters X , Y , and Z , possibly with subscripts, to denote variables, but for the measures we continue to use the propositional version. Also, in these examples Δ will consist of atoms to illustrate a method for assigning weights.

Example 30. Let $\Delta = \{\text{salary}(\text{bob}, 1000), \text{salary}(\text{bob}, 2000)\}$ and consider the following weights where we assume that the weight is dependent on the range of values for the salary for Bob.

	w
$\text{salary}(\text{bob}, 1000)$	10
$\text{salary}(\text{bob}, 2000)$	10

Let Γ be the integrity constraints obtained by instantiating the variables X_1 and X_2 in the following scheme with terms in the language of the proper data type such that $X_1 \neq X_2$

$$\text{salary}(\text{bob}, X_1) \rightarrow \neg \text{salary}(\text{bob}, X_2)$$

So, in this case, Γ consists of the following integrity constraints:

$$\begin{aligned} \text{salary}(\text{bob}, 1000) &\rightarrow \neg \text{salary}(\text{bob}, 2000) \\ \text{salary}(\text{bob}, 2000) &\rightarrow \neg \text{salary}(\text{bob}, 1000) \end{aligned}$$

Using the Manhattan measure with these weights we obtain

$$\text{Profiles}_d^{\min}(\Delta, \Gamma) = \{(10, 0), (0, 10)\}.$$

Therefore, the inconsistency measures for the weighted measure obtained by using the above weights are $I_d^{\text{sum}}(\Delta) = 10$, $I_d^{\text{max}}(\Delta) = 10$, and $I_d^{\text{hit}}(\Delta) = 1$.

We have suggested that the weights could be chosen so that the significance of the inconsistency rises as the difference in the values taken by the data deviate. In order to assign the weights, we may choose to use an equation, as we illustrate in the following example where we consider weight to be a linear function of the difference between the given value and the median value.

Example 31. Let $\Delta = \{\text{salary}(\text{bob}, 1000), \text{salary}(\text{bob}, 1100), \text{salary}(\text{bob}, 1400), \text{salary}(\text{bob}, 1500), \text{salary}(\text{bob}, 1600), \text{salary}(\text{bob}, 1900), \text{salary}(\text{bob}, 2000)\}$. Consider the following weights where we assume that the weight is dependent on the range of values for the salary for Bob. So the most extreme values for the salary (i.e. 1000 and 2000) have highest significance, whereas the least extreme value (i.e. 1500) has the lowest significance. We capture this by the following equation where X^* is the mid-point between the minimum and maximum value for the salary and 1 is added to avoid null weights.

$$w(\text{salary}(\text{bob}, X)) = \frac{|X - X^*|}{100} + 1$$

Using this equation, we get the following weight for the example.

	w
$\text{salary}(\text{bob}, 1000)$	6
$\text{salary}(\text{bob}, 1100)$	5
$\text{salary}(\text{bob}, 1400)$	2
$\text{salary}(\text{bob}, 1500)$	1
$\text{salary}(\text{bob}, 1600)$	2
$\text{salary}(\text{bob}, 1900)$	5
$\text{salary}(\text{bob}, 2000)$	6

Also, suppose that Γ is the set of integrity constraints obtained from the scheme $\text{salary}(\text{bob}, X_1) \rightarrow \neg \text{salary}(\text{bob}, X_2)$ where $X_1 \neq X_2$ (same as for Example 30). In this case, $\text{Profiles}_d^{\min}(\Delta, \Gamma)$ consists of seven 7-tuples obtained from $(6, 5, 2, 1, 2, 5, 6)$ by changing one of the numbers to 0. For example, $(6, 5, 2, 1, 2, 5, 0) \in \text{Profiles}_d^{\min}(\Delta, \Gamma)$. Again, using the Manhattan measure with these weights we obtain $I_d^{\text{sum}}(\Delta) = 21$, $I_d^{\text{max}}(\Delta) = 6$, and $I_d^{\text{hit}}(\Delta) = 6$.

Finally, we consider an example with two dimensions in the inconsistency of the data. Here the data can be inconsistent with the integrity constraints with respect to just age or just salary or both age and salary.

Example 32. Let Δ be the set containing the following atoms.

$$\begin{array}{lll} \text{salary}(\text{bob}, 40, 1000) & \text{salary}(\text{bob}, 50, 1000) & \text{salary}(\text{bob}, 60, 1000) \\ \text{salary}(\text{bob}, 40, 1200) & \text{salary}(\text{bob}, 50, 1200) & \text{salary}(\text{bob}, 60, 1200) \\ \text{salary}(\text{bob}, 40, 1400) & \text{salary}(\text{bob}, 50, 1400) & \text{salary}(\text{bob}, 60, 1400) \end{array}$$

α	T	T	T	B	B	B	F	F	F
β	T	B	F	T	B	F	T	B	F
$\alpha \vee \beta$	T	T	T	T	B	B	T	B	F
$\alpha \wedge \beta$	T	B	F	B	B	F	F	F	F
$\neg\alpha$	F	F	F	B	B	B	T	T	T

Figure 2: Truth table for the three valued logic (3VL). This semantics extends the classical semantics with a third truth value, B , denoting “contradictory”. Columns 1, 3, 7, and 9 are the classical semantics; the other columns give the extended semantics.

Consider the following weights where we assume that the weight is dependent on the range of values for the age and salary for Bob. So, the most extreme values for the salary (i.e. 1000 and 1400) have highest significance, whereas the least extreme value (i.e. 1200) has the lowest significance. Similarly, the most extreme values for the age (i.e. 40 and 60) have highest significance, whereas the least extreme value (i.e. 50) has the lowest significance. We capture this by the following equation where X^* is the mid-point between the minimum and maximum value for the age and Y^* is the mid-point between the minimum and maximum value for the salary, and 1 is added to avoid null weights.

$$w(\text{salary}(\text{bob}, X, Y)) = \frac{|X - X^*|}{10} + \frac{|Y - Y^*|}{100} + 1$$

Using this equation, we get the following weights for this example.

	w		w		w
$\text{salary}(\text{bob}, 40, 1000)$	4	$\text{salary}(\text{bob}, 50, 1000)$	3	$\text{salary}(\text{bob}, 60, 1000)$	4
$\text{salary}(\text{bob}, 40, 1200)$	2	$\text{salary}(\text{bob}, 50, 1200)$	1	$\text{salary}(\text{bob}, 60, 1200)$	2
$\text{salary}(\text{bob}, 40, 1400)$	4	$\text{salary}(\text{bob}, 50, 1400)$	3	$\text{salary}(\text{bob}, 60, 1400)$	4

Also, suppose that Γ is composed of the integrity constraints obtained from the scheme $\text{salary}(\text{bob}, X_1, Y_1) \rightarrow \neg\text{salary}(\text{bob}, X_2, Y_2)$ where $X_1 \neq X_2$ or $Y_1 \neq Y_2$ as in the previous examples. In this case, $\text{Profiles}_d^{\min}(\Delta, \Gamma)$ consists of 9 9-tuples obtained from $(4, 2, 4, 3, 1, 3, 4, 2, 4)$ by changing one of the numbers to 0. For example, $(4, 2, 4, 3, 1, 3, 0, 2, 4) \in \text{Profiles}_d^{\min}(\Delta, \Gamma)$. Again, using the Manhattan measure with these weights we obtain $I_d^{\text{sum}}(\Delta) = 23$, $I_d^{\text{max}}(\Delta) = 4$, and $I_d^{\text{hit}}(\Delta) = 8$.

Taking significance into account using these measures means that we consider how “incorrect” or how extreme the literals are. Smaller ranges of values in the data have lower weights than wider ranges of values in the data.

9 Comparison with existing measures

As we stated in Section 1, there are a number of existing proposals for measures of inconsistency. We review these in this section, and compare them with our new measures. For this, we provide here some definitions taken from [GH11] where these measures are studied in detail.

For a knowledgebase Δ , $\text{MI}(\Delta)$ is the set of minimal inconsistent subsets of Δ , and $\text{MC}(\Delta)$ is the set of maximal consistent subsets of Δ . Also, if $\text{MI}(\Delta) = \{M_1, \dots, M_n\}$ then $\text{Problematic}(\Delta) = M_1 \cup \dots \cup M_n$, and $\text{Free}(\Delta) = \Delta \setminus \text{Problematic}(\Delta)$. Thus $\text{Problematic}(\Delta)$ contains the formulae in Δ that are involved in at least one inconsistency, and the formulae of $\text{Free}(\Delta)$ are not involved in any inconsistency. The set of formulae in Δ that are individually inconsistent is given by $\text{Selfcontradictions}(\Delta)$.

One approach uses Priest’s three valued logic (3VL) [Pri79] with the classical two valued semantics augmented by a third truth value denoting inconsistency. The truth values for the connectives are defined in Figure 2. An interpretation i is a function that assigns to each atom that appears in Δ one of three truth values: $i : \text{Atoms}(\Delta) \rightarrow \{F, B, T\}$. For an interpretation i it is convenient to separate the atoms into two groups, namely the ones that are assigned a classical

truth value and the ones that are assigned B as follows: $\text{Binarybase}(i) = \{\alpha \mid i(\alpha) = T \text{ or } i(\alpha) = F\}$; and $\text{Conflictbase}(i) = \{\alpha \mid i(\alpha) = B\}$. For a knowledgebase Δ we define the bi-models as the set of interpretations where no formula in Δ is assigned the truth value F : $\text{BiModels}(\Delta) = \{i \mid \text{for all } \phi \in \Delta, i(\phi) = T \text{ or } i(\phi) = B\}$. Then, as a measure of inconsistency for Δ we define $\text{Contension}(\Delta) = \text{Min}\{|\text{Conflictbase}(i)| \mid i \in \text{BiModels}(\Delta)\}$. This is the minimal number of atoms that need to be assigned B in order to get a 3VL model of Δ .

Example 33. For $\Delta = \{a, \neg a, a \vee b, \neg b\}$, there are two models of Δ , i_1 and i_2 , where $i_1(a) = B$, $i_1(b) = B$, and $i_2(a) = B$, $i_2(b) = F$. Therefore, $|\text{Conflictbase}(i_1)| = 2$ and $|\text{Conflictbase}(i_2)| = 1$. Hence, $\text{Contension}(\Delta) = 1$.

Next, we define five inconsistency measures from the literature: $I_C(\Delta)$ counts the number of minimal inconsistent subsets of Δ ; $I_M(\Delta)$ counts the sum of the number of maximal consistent subsets together with the number of contradictory formulae but 1 must be subtracted to make $I(\Delta) = 0$ when Δ is consistent; $I_P(\Delta)$ counts the number of formulae in minimal inconsistent subsets of Δ ; $I_B(\Delta)$ counts the minimum number of atoms that need to be assigned B amongst the 3VL models of Δ ; and I_Q computes the weighted sum of the minimal inconsistent subsets of Δ , where the weight is the inverse of the size of the minimal inconsistent subset (and hence smaller minimal inconsistent subsets are regarded as more inconsistent than larger ones).

Definition 28. For a knowledgebase Δ , the inconsistency measures I_C , I_P , I_B , I_S , and I_R are:

- $I_C(\Delta) = |\text{MI}(\Delta)|$
- $I_M(\Delta) = (|\text{MC}(\Delta)| + |\text{Selfcontradictions}(\Delta)|) - 1$
- $I_P(\Delta) = |\text{Problematic}(\Delta)|$
- $I_B(\Delta) = \text{Contension}(\Delta)$
- $I_Q(\Delta) = \begin{cases} 0 & \text{if } \Delta \text{ is consistent} \\ \sum_{X \in \text{MI}(\Delta)} \frac{1}{|X|} & \text{otherwise} \end{cases}$

In [GH11], we showed that the five measures given above are pairwise order-incompatible. In Proposition 5 we showed that the three new inconsistency measures defined in this paper are also pairwise order-incompatible. Now we can show the pairwise order-incompatibility of all eight inconsistency measures. But first we show an interesting connection between I_B and I_D^{sum} for the Dalal measure d .

Proposition 17. For the Dalal distance measure D , for all Δ , $I_B(\Delta) \leq I_D^{sum}(\Delta)$.

Proof. Let $x = I_D^{sum}(\Delta)$. Then $x = \sum_{i=1}^n k_i$ where $(k_1, \dots, k_n) \in \text{Profiles}_D^{min}(\Delta)$. Each k_i represents changing the truth values of k_i atoms, which means that the assignment of the truth value B to x (not necessarily distinct) atoms provides a 3VL model of Δ . Hence, $I_B(\Delta) \leq I_D^{sum}(\Delta)$. \square

We now give an example where $I_B(\Delta) < I_D^{sum}(\Delta)$.

Example 34. Let $\Delta_5 = \{a \wedge b, a \wedge c, \neg a \wedge b, \neg a \wedge c\}$. Here $I_B(\Delta) = 1 < I_D^{sum}(\Delta) = 2$.

Assuming the five measures I_C , I_M , I_P , I_B , I_Q are pairwise order-incompatible (as shown in [GH11]), and I_d^{sum} , I_d^{max} , and I_d^{hit} are all pairwise order-incompatible (as shown in Proposition 5), we now show that each of I_d^{sum} , I_d^{max} , and I_d^{hit} is pairwise order-incompatible with each of I_C , I_M , I_P , I_B , I_Q .

Theorem 3. The eight inconsistency measures: I_C , I_M , I_P , I_B , I_Q , I_d^{sum} , I_d^{max} , and I_d^{hit} are all pairwise order-incompatible.

Proof. By [GH11] and Proposition 5 it suffices to show that each of $\{I_C, I_M, I_P, I_B, I_Q\}$ is order-incompatible with each of $\{I_d^{sum}, I_d^{max}, I_d^{hit}\}$. We use four knowledgebases from Examples 8, 9, 11, and 34. The following table shows the various inconsistency measures for these knowledgebases.

Δ	I_D^{sum}	I_D^{max}	I_D^{hit}	I_C	I_M	I_P	I_B	I_Q
Δ_1	2	1	1	1	1	2	2	1/2
Δ_2	1	1	1	1	2	3	1	1/3
Δ_4	3	1	3	3	7	6	3	3/2
Δ_5	2	1	2	4	1	4	1	2

We therefore get the following differences between each of I_D^{sum} , I_D^{max} , and I_D^{hit} , with each of I_C , I_M , I_P , I_B , and I_Q .

	I_D^{sum}	I_D^{max}	I_D^{hit}
I_C	$I_D^{sum}(\Delta_1) > I_D^{sum}(\Delta_2)$ but $I_C(\Delta_1) = I_C(\Delta_2)$	$I_D^{max}(\Delta_1) = I_D^{max}(\Delta_5)$ but $I_C(\Delta_1) < I_C(\Delta_5)$	$I_D^{hit}(\Delta_4) > I_D^{hit}(\Delta_5)$ but $I_C(\Delta_4) < I_C(\Delta_5)$
I_M	$I_D^{sum}(\Delta_1) > I_D^{sum}(\Delta_2)$ but $I_M(\Delta_1) < I_M(\Delta_2)$	$I_D^{max}(\Delta_1) = I_D^{max}(\Delta_2)$ but $I_M(\Delta_1) < I_M(\Delta_2)$	$I_D^{hit}(\Delta_1) = I_D^{hit}(\Delta_2)$ but $I_M(\Delta_1) < I_M(\Delta_2)$
I_P	$I_D^{sum}(\Delta_1) > I_D^{sum}(\Delta_2)$ but $I_P(\Delta_1) < I_P(\Delta_2)$	$I_D^{max}(\Delta_1) = I_D^{max}(\Delta_2)$ but $I_P(\Delta_1) < I_P(\Delta_2)$	$I_D^{hit}(\Delta_1) = I_D^{hit}(\Delta_2)$ but $I_P(\Delta_1) < I_P(\Delta_2)$
I_B	$I_D^{sum}(\Delta_1) = I_D^{sum}(\Delta_5)$ but $I_B(\Delta_1) > I_B(\Delta_5)$	$I_D^{max}(\Delta_1) = I_D^{max}(\Delta_2)$ but $I_B(\Delta_1) > I_B(\Delta_2)$	$I_D^{hit}(\Delta_1) = I_D^{hit}(\Delta_2)$ but $I_B(\Delta_1) > I_B(\Delta_2)$
I_Q	$I_D^{sum}(\Delta_1) = I_D^{sum}(\Delta_5)$ but $I_Q(\Delta_1) < I_Q(\Delta_5)$	$I_D^{max}(\Delta_1) = I_D^{max}(\Delta_2)$ but $I_Q(\Delta_1) > I_Q(\Delta_2)$	$I_D^{hit}(\Delta_1) = I_D^{hit}(\Delta_2)$ but $I_Q(\Delta_1) > I_Q(\Delta_2)$

□

Whilst we have not directly compared our distance-based measures to some more recent proposals for syntax-based measures (e.g. vectorial measures [MLJ11], deduction-based measures [JR13], and measures based on maximum consistent subsets [RSB15]), we believe that the above results will extend in a straightforward way. Similarly, we believe that it is straightforward to extend the above results to see the difference with the proposals that give a higher value for inconsistency to smaller minimal inconsistent sets of formulae (e.g. [Kni01, Kni03, DRMO10]).

Measures of inconsistency have also been developed for description logics (e.g. [MQHL07, QH07, ZHQ⁺09, MH10]). Ma *et al* propose a distance-based measure of inconsistency and incoherence [MH10] which considers the distance between each formula and an interpretation, and offers an analogous definition to our d-sum inconsistency measure. The paper then investigates a normalized measure based on a notion of deviation.

Other kinds of measures of inconsistency have been developed for probabilistic knowledge [Thi13, Pot14, DF15] and fuzzy knowledge [Mui11]. These augment logical information with quantitative information, and the measures consider distance with respect to numerical assignments. Therefore, we will not consider those further here.

In the literature on measures of inconsistency, there are also proposals for ascribing the blame for inconsistencies to individual formulae [HK10, MLJ12, Mu15] or to sources [CPRT15]. [Mu15] is based on the causal model, while the others use measures of inconsistency for the knowledgebase to distribute the blame using the Shapley value. We do not investigate this further here, but we could use our distance-based measures together with the notion of a Shapley value to ascribe blame to individual formulae.

Taking the significance of inconsistency into account when measuring inconsistency was first raised in [Hun03]. This included a proposal for extending a measure of inconsistency based on four-valued models to a many-valued logic. So instead of a single value for “both” (i.e. the value for both true and false), there are many values for “both”. These are linearly ordered from “least significant” to “most significant”. Given that the underlying measure of inconsistency is based on four-valued logic, there are substantial differences with distance-based measures (as indicated by the above result concerning the I_B measure). Furthermore, the notion of significance appears more general in the distance-based framework, and perhaps it is more transparent. Nonetheless, in future work, it would be interesting to more fully investigate the relationship between the two approaches.

Significance has also been considered in a framework for measuring inconsistency based on analyzing inconsistent formulae in terms of minimally inconsistent subsets [MJLL05, MJL⁺13] where the significance takes account of the relative priority (i.e. importance) of each formula in the set. We believe that the above results extend in a straightforward way to show the difference between this syntax-based approach and our model-theoretic approach.

The main conclusion to draw from this comparison is that distance-based measures offer a new set of alternatives for measuring inconsistency. These provide further insights into the nature of inconsistency, and they may prove to be useful in specific applications.

10 Discussion

This paper makes the following contributions: (1) A general framework for measuring inconsistency based on semantic distance between models of the formulae in a knowledgebase; (2) An investigation of semantic distance based on distance measures such as Manhattan distance and Euclidean distance; (3) An investigation into the use of weighted distance, non-weighted distance, cost functions, and cost rankings, to capture a notion of significance of inconsistency; and (4) Applications of the new measures in analyzing inconsistent databases. We have also shown how this new class of measures offers a novel alternative to existing proposals, and provides further insights into the nature of inconsistency.

Returning to the original idea of distance-based measures, we saw a natural correspondence with the distance-based merging operators proposed by Konieczny and Pino-Perez [KP98], and the idea of measuring inconsistency using distance. We see this correspondence between the Max merging operator and the d-max inconsistency measure, and the Sum merging operator and the d-sum inconsistency measure. It may therefore be interesting to consider whether there is an inconsistency measure that corresponds to the Gmax merging operator. The challenge is that translating the idea of Gmax would call for an inconsistency measure that is a tuple of values rather than an individual value. This may be appropriate for some applications, but it would take us away from the usual postulates for an inconsistency measure which we have assumed for this paper.

An alternative approach to harnessing Gmax may be to treat each tuple in $\text{Profiles}_d^{\text{min}}(\Delta)$ as a number. For a tuple $(k_1, \dots, k_n) \in \text{Profiles}_d^{\text{min}}(\Delta)$, $x = \text{Max}\{k_1, \dots, k_n\}$ gives the biggest number x in the tuple. Then we can treat the tuple of n numbers (base 10) as a single number to the base $x + 1$. Before we do this, we should also reorder the numbers in the tuple from largest to smallest. For this, we introduce the function $\text{DownOrder}((k_1, \dots, k_n))$ which returns the tuple (q_1, \dots, q_n) where $q_1 \geq q_2 \geq \dots \geq q_n$. For example, $\text{DownOrder}((0, 1, 2, 1, 2)) = (2, 2, 1, 1, 0)$ and $\text{DownOrder}((0, 2, 2, 0, 2)) = (2, 2, 2, 0, 0)$. Then we use the function Base_{x+1} to return the base $x + 1$ number. For example, $\text{Base}_{x+1}((2, 2, 1, 1, 0)) = 22110_3 = 228$, while $\text{Base}_{x+1}((2, 2, 2, 0, 0)) = 22200_3 = 234$.

Definition 29. Let $\Delta \subseteq \mathcal{L}$ be a set of propositional formulae where each $\phi_i \in \Delta$ is consistent, and let d be a distance measure. The **d-gmax inconsistency measure** is defined as follows.

$$I_d^{\text{gmax}}(\Delta) = \text{Min}\{(q_1 \dots q_n)_{x+1} \mid (k_1, \dots, k_n) \in \text{Profiles}_d^{\text{min}}(\Delta) \text{ and } x = \max(\{k_1, \dots, k_n\}) \\ \text{and } \text{DownOrder}((k_1, \dots, k_n)) = (q_1, \dots, q_n)\}$$

The aim of the gmax inconsistency measure is to find a compromise that is more evenly distributed across the formulae. Furthermore, the number returned gives more information about the nature of that distribution. We illustrate this next.

Example 35. Let $\Delta = \{a \wedge b, c \wedge d, \neg a \wedge \neg b \wedge \neg c \wedge \neg d\}$. Using the Dalal distance D , $\text{Profiles}_D^{\text{min}}(\Delta) = \{(2, 2, 0), (2, 1, 1), (1, 2, 1), (1, 0, 3), (0, 1, 3), (0, 0, 4)\}$. Hence, $I_D^{\text{sum}}(\Delta) = 4$, $I_D^{\text{max}}(\Delta) = 2$, $I_D^{\text{hit}}(\Delta) = 1$, and $I_D^{\text{gmax}}(\Delta) = 211_3 = 22$.

We leave further development of the gmax inconsistency measure to future work. We also leave the development of algorithms for calculating distance-based measures to future work. Nonethe-

less, we believe that the correspondence between distance-based merging operators and distance-based inconsistency measures means that this class of inconsistency measures can be viably calculated using binary decision diagrams [GH08].

Also in future work, we will investigate further instances of distance-based measures of inconsistency, develop potential applications of distance-based measures (e.g. context-sensitive approaches to dealing with inconsistency such as considered in [SA07]), develop efficient ways of getting weights, develop efficient ways of defining non-weighted distance measures, and consider Besnard’s framework of postulates for inconsistency measures [Bes14] as an alternative to the postulates we used in this paper. We also plan to address some of the shortcomings of using the Hamming distance, as discussed by Lafage and Lang [LL01], by using distances based on Choquet integrals.

Acknowledgement We thank the reviewers for many excellent comments. We also thank Glauber De Bona for valuable feedback.

References

- [Bes14] Ph. Besnard. Revisiting postulates for inconsistency measures. In *Logics in Artificial Intelligence (JELIA ’14)*, LNCS, pages 383–396. Springer, 2014.
- [BL02] I. Bloch and J. Lang. Towards Mathematical Morpho-Logics. In *Technologies for Constructing Intelligent Systems*, volume 2, pages 367–380. Springer-Verlag, 2002.
- [CPRT15] L. Cholvy, L. Perrussel, W. Raynaut, and J.-M. Thévenin. Consistency-based reliability assessment. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU’15)*, volume 9161 of LNCS, pages 246–256. Springer, 2015.
- [Dal88] M. Dalal. Investigations into a theory of knowledge base revision. In *Proceedings of the Seventh National Conference on Artificial Intelligence, AAAI’88*, volume 2, pages 475–479, 1988.
- [DF15] D. De Bona and M. Finger. Measuring inconsistency in probabilistic logic: rationality postulates and Dutch book interpretation. *Artificial Intelligence*, 277:140–164, 2015.
- [DRMO10] D. Doder, M. Raskovic, Z. Markovic, and Z. Ognjanovic. Measures of inconsistency and defaults. *International Journal of Approximate Reasoning*, 51(7):832–845, 2010.
- [GH08] N. Gorogiannis and A. Hunter. Implementing semantic merging operators using binary decision diagrams. *International Journal of Approximate Reasoning*, 49(1):234–251., 2008.
- [GH11] J. Grant and A. Hunter. Measuring consistency gain and information loss in stepwise inconsistency resolution. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU’11)*, volume 6717 of LNCS, pages 362–373. Springer, 2011.
- [GH13] J. Grant and A. Hunter. Distance-based measures of inconsistency. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU’13)*, volume 7958 of LNCS, pages 230–241. Springer, 2013.
- [Gra78] J. Grant. Classifications for inconsistent theories. *Notre Dame Journal of Formal Logic*, 19:435–444, 1978.
- [HK06] A. Hunter and S. Konieczny. Shapley inconsistency values. In *Proceedings of the 10th International Conference on Knowledge Representation (KR’06)*, pages 249–259. AAAI Press, 2006.

- [HK10] A. Hunter and S. Konieczny. On the measure of conflicts: Shapley inconsistency values. *Artificial Intelligence*, 174:1007–1026, 2010.
- [Hun03] A. Hunter. Evaluating the significance of inconsistencies. In *Proceedings of the International Joint Conference on AI (IJCAI'03)*, pages 468–473. Morgan Kaufmann, 2003.
- [JR13] S. Jabbour and B. Raddaoui. Measuring inconsistency through minimal proofs. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'13)*, LNCS, pages 290–301. Springer, 2013.
- [Kni01] K. Knight. Measuring inconsistency. *Journal of Philosophical Logic*, 31:77–98, 2001.
- [Kni03] K. Knight. Two information measures for inconsistent sets. *Journal of Logic, Language and Information*, 12:227–248, 2003.
- [KP98] S. Konieczny and R. Pino Pérez. On the logic of merging. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR'98)*, pages 488–498, 1998.
- [LL01] C. Lafage and J. Lang. Propositional distances and preference representation. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'01)*, volume 2143 of LNCS. Springer, 2001.
- [MH10] Y. Ma and P. Hitzler. Distance-based measures of inconsistency and incoherency for description logic. In *Proceedings of the International Workshop on Description Logics (DL'10)*, volume 573 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2010.
- [MJL⁺13] K. Mu, Z. Jin, W. Liu, D. Zowghi, and B. Wei. Measuring the significance of inconsistency in the viewpoints framework. *Science of Computer Programming*, 78(9):1572–1599, 2013.
- [MJLL05] K. Mu, Z. Jin, R. Lu, and W. Liu. Measuring inconsistency in requirements engineering. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'05)*, volume 3571 of LNAI, pages 440–451. Springer, 2005.
- [MLJ11] K. Mu, W. Liu, and Z. Jin. A general framework for measuring inconsistency through minimal inconsistent sets. *Knowledge and Information Systems*, 27:85–114, 2011.
- [MLJB11] K. Mu, W. Liu, Z. Jin, and D. Bell. A Syntax-based approach to measuring the degree of inconsistency for belief base. *International Journal of Approximate Reasoning*, 52(7):978–999, 2011.
- [MLJ12] K. Mu, W. Liu, and Z. Jin. Measuring the blame of each formula for inconsistent prioritized knowledge bases. *Journal of Logic and Computation*, 22(3):481–516, 2012.
- [MQHL07] Y. Ma, G. Qi, P. Hitzler, and Z. Lin. Measuring inconsistency for description logics based on paraconsistent semantics. In *Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'07)*, volume 4724 of LNCS, pages 30–41. Springer, 2007.
- [Mu15] K. Mu. Responsibility for inconsistency. *International Journal of Approximate Reasoning*, 61:43–60, 2015.
- [Mui11] D. Muiño. Measuring and repairing inconsistency in knowledge bases with graded truth. *Fuzzy Sets and Systems*, 197:108–122, 2011.
- [Pot14] N. Potyka. Linear programs for measuring inconsistency in probabilistic logics. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR'14)*, 2014.

- [Pri79] G. Priest. Logic of paradox. *Journal of Philosophical Logic*, 8:219–241, 1979.
- [QH07] G. Qi and A. Hunter. Measuring incoherence in description logic-based ontologies. In *Proceedings of the International Semantic Web Conference (ISWC'07)*, volume 4825 of *LNCS*, pages 381–394. Springer, 2007.
- [RSB15] B. Raddaoui, Y. Salhi, and O. Brahim. On measuring inconsistency using maximal consistent sets. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'15)*, LNCS. Springer, 2015.
- [SA07] V. S. Subrahmanian and L. Amgoud. A general framework for reasoning about inconsistency. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI'07)*, pages 599–504, 2007.
- [Thi13] M. Thimm. Inconsistency measures for probabilistic logics. *Artificial Intelligence*, 197:1–24, 2013.
- [ZHQ⁺09] L. Zhou, H. Huang, G. Qi, Y. Ma, Z. Huang, and Y. Qu. Measuring inconsistency in DL-Lite ontologies. In *Proceedings of the IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, 2009*, pages 349–356. IEEE, 2009.