Approaches to Measuring Inconsistent Information

Anthony Hunter[1] and Sébastien Konieczny[2]

[1] Department of Computer Science,
University College London, Gower Street,
London WC1E 6BT, UK
a.hunter@cs.ucl.ac.uk

[2] CRIL-CNRS, Université d'Artois, 62300 Lens, France
konieczny@cril.univ-artois.fr

Abstract. Measures of quantity of information have been studied extensively for more than fifty years. The seminal work on information theory is by Shannon [67]. This work, based on probability theory, can be used in a logical setting when the worlds are the possible events. This work is also the basis of Lozinskii's work [48] for defining the quantity of information of a formula (or knowledgebase) in propositional logic. But this definition is not suitable when the knowledgebase is inconsistent. In this case, it has no classical model, so we have no "event" to count. This is a shortcoming since in practical applications (e.g. databases) it often happens that the knowledgebase is not consistent. And it is definitely not true that all inconsistent knowledgebases contain the same (null) amount of information, as given by the "classical information theory". As explored for several years in the paraconsistent logic community, two inconsistent knowledgebases can lead to very different conclusions, showing that they do not convey the same information. There has been some recent interest in this issue, with some interesting proposals. Though a general approach for information theory in (possibly inconsistent) logical knowledgebases is missing. Another related measure is the measure of contradiction. It is usual in classical logic to use a binary measure of contradiction: a knowledgebase is either consistent or inconsistent. This dichotomy is obvious when the only deductive tool is classical inference, since inconsistent knowledgebases are of no use. But there are now a number of logics developed to draw non-trivial conclusions from an inconsistent knowledgebase. So this dichotomy is not sufficient to describe the amount of contradiction of a knowledgebase, one needs more fine-grained measures. Some interesting proposals have been made for this. The main aim of this paper is to review the measures of information and contradiction, and to study some potential practical applications. This has significant potential in developing intelligent systems that can be tolerant to inconsistencies when reasoning with real-world knowledge.

L. Bertossi et al. (Eds.): Inconsistency Tolerance, LNCS 3300, pp. 191–236, 2004.

1 Introduction

Traditionally the consensus of opinion in the computer science community is that inconsistency is undesirable. Many believe that databases, knowledgebases, and software specifications, should be completely free of inconsistency, and try to eradicate inconsistency from them by any means possible. Others address inconsistency by isolating it, and perhaps resolving it locally. All seem to agree, however, that data of the form q and $\neg q$, for any proposition q cannot exist together, and that the conflict must be resolved somehow.

This view is too simplistic for developing robust intelligent systems, and furthermore, it fails to use the benefits of inconsistency in intelligent activities. Inconsistency in information is the norm in the real-world, and so should be formalized and used, rather than always rejected [23].

There are cases where q and $\neg q$ can be perfectly acceptable together and hence need not be resolved. Consider for example an income tax database where contradictory information on a taxpayer can be useful evidence in a fraud investigation. Maybe the taxpayer has completed one form that states the taxpayer has 6 children (and hence get the tax benefits for that) and completed another that states the taxpayer has 0 children. In other cases, q and $\neg q$ serve as a useful trigger for various logical actions. Inconsistency is useful in directing reasoning, and instigating the natural processes of argumentation, information seeking, multi-agent interaction, knowledge acquisition and refinement, adaptation, and learning.

Of course, there are inconsistencies that do need to be resolved. But, the decision to resolve, and the approach to resolution, needs to be context-sensitive. There is also the question of when to resolve inconsistencies. Immediate resolution of inconsistencies can result in the loss of valuable information if an arbitrary choice is made on what to reject. Consider for example the requirements capture stage in software engineering. Here premature resolution can force an arbitary decision to be made without the choice being properly considered. This can therefore overly constrain the requirements capture process.

Similarly when working with distributed databases, it cannot be expected that there are no conflicts between the databases. Conflicts in this case can have different meanings. It can sometimes denote an error in some database, in which case we can simply use a database repair. But more often conflicts will denote deeper disagreement between sets of databases, with no easy repair. So, in this case, resolution of all conflicts is not the solution, since we need to keep track of the conflict. The straighforward reason is that "having no information about some fact" or "having contradictory information about some fact" cannot be regarded as having the same epistemic status. After a repair of a set of databases, either we forget all information about the facts in conflict, or we decide what is the correct answer (among the conflicting ones). But, for the user (human or software), it is not the same thing to receive an answer "the fact A is true" or "the fact A seems to be true, but there is a conflict about it". Such answers, needed in high-level reasoning systems, require us to not resolve the conflicts (see for example [12]).

The call for robust, and intelligent, systems, has led to an increased interest in inconsistency tolerance in computer science. The central position is that the collapse of classical logic in cases of inconsistency should be circumvented. In other words, we need to suspend the axiom of absurdity (*ex falso quodlibet*) for many kinds of reasoning. A number of useful proposals have been made in the field of paraconsistent logics (see for example [28, 13]).

In addition, we need strategies for analysing inconsistent information. This need has in part driven the approach of argumentation systems which compare pros and cons for potential conclusions from conflicting information (for reviews see [58, 14]). Also important are strategies for isolating inconsistency and for taking appropriate actions, including resolution actions. This calls for uncertainty reasoning and meta-level reasoning. Furthermore, the cognitive activities involved in reasoning with inconsistent information need to be directly related to the kind of inconsistency. So, in general, we see the need for inconsistency tolerance giving rise to a range of technologies for inconsistency management. These in turn call for richer ways of describing and comparing conflicts.

Comparing heterogeneous sources often involves comparing conflicts. Suppose we are dealing with a group of clinicians advising on some patient, a group of witnesses of some incident, or a set of newspaper reports covering some event. These are all situations where we expect some degree of inconsistency in the information. Suppose that the information by each source i is represented by the set Φ_i. Each source may provide information that conflicts with the domain knowledge Ψ. Let us represent $\Phi_i \cup \Psi$ by Δ_i for each source i. Now, we may want to know whether one source is more inconsistent than another — so whether Δ_i is more inconsistent that Δ_j — and in particular determine which is the least inconsistent of the sources and so identify a minimal Δ_i in this inconsistency ordering. We may then view this minimal knowledgebase as the least problematical or most reliable source of information. This point is close to the notion of verisimilitude, as initiated by Popper [57, 44, 63].

When an autonomous system works with a set of information, beliefs, knowledge, preferences, ... expressed in a logical form (we will talk about pieces of information in the following instead of always specifying information, belief, knowledge, preferences), the notion of informational content of a piece of information and the notion of amount of contradiction are of crucial interest. Effectively, in many high-level reasoning tasks one needs to know what is the amount of information conveyed by a piece of information and/or what is the amount of contradiction involved with this piece of information. This is particularly important in complex information about the real world where inconsistencies are hard to avoid.

While information measures enable us to say how "valuable" a piece of information is by showing how precise it is, contradiction measures enable us to say how "unvaluable" a piece of information is by showing how conflicting it is. As joint/conditional information measures are useful to define a notion of pertinence of a new piece of information with respect to an old one (or more generally for a set of information), joint/conditional contradiction measures can be useful to define a notion of conflict between pieces of information, that can be useful for many applications. These two measures are to a large extent independent of one another, but needed in numerous applications, for instance:

- In diagnosis, some initial assumptions stating that each component works normally are made; those assumptions may conflict with actual observations. Measuring the conflict of the resulting base may be a good indication about how hard it will be to identify the faulty components.
- In belief revision, when an agent receives a new piece of information which contradicts her previous beliefs, evaluating how much this information is conflicting with the previous beliefs can be useful to decide whether the agent accepts or rejects the new piece of information.
- In belief merging, degrees of information and contradiction can be the basis on which one can decide whether to take account or not of the information being conveyed by an agent. If the degree of contradiction of the information given by an agent is high, it may be relevant to reject the information, since there is some significant evidence that the source is not reliable; however, this must be balanced by the quantity of information furnished by the agent, especially when she also gives some important and uncontroversial pieces of information.

One of the applications discussed above concerns the problem of iterated belief revision. The problem of belief revision is to incorporate a new piece of information which is more reliable than (and conflicting with) the old beliefs of the agent. This problem has received a nice answer in the work of Alchourron, Gardenfors, Makinson [1] in the one-step case. But when one wants to iterate revision (i.e. to generalize it to the n-step case), there are numerous problems and no definitive answer has been reached in the purely qualitative case [16, 22]. Using a partially quantitative framework, some proposals have given interesting results (see e.g. [69, 68]). Here "partially quantitative" means that the incoming piece of information needs to be labelled by a degree of confidence denoting how strongly we believe it. The problem in this framework is to justify the use of such a degree, what does it mean exactly and where does it come from. So if one can define composite measures, from the information measure and the contradiction measure, then one can define several policies for the agent (we can figure out an agent who accepts a new piece of information only if it brings more information than contradiction, etc). We can then use the "partially quantitative" framework to derive revision operators with a nice behaviour. In this setting, since the degree attached to the incoming information is not a given data, but computed directly from the incoming information and the agent policy (behaviour with respect to information and contradiction, encoded by a composite measure) then the problem of the justification of the meaning of the degrees is avoided.

Another related application is the use of degrees of conflict and information to the problem of belief merging. Given a set of agents with conflicting beliefs, the problem of belief merging is to know how to define the beliefs of the group. A natural way to define the result of the merging is to see the group as a set of agents involved in a game (this can be intuitively explained as a modelisation of a human meeting), and look for winning coalitions of agents. An example of a definition of coalition can be a set of agent with consistent beliefs (or minimal conflicting ones) and a maximal joint degree of information. Then for determining the winning coalition we can look at the degree of conflict and define the winning coalition as the one which is minimally conflicting with the others. Other interesting strategies can be defined as well.

These two examples show that the conjoint use of degree of information and contradiction can open a huge scope of research. The two given examples are actually original approaches to revision and fusion. Similar examples can be found for other reasoning tasks. This highlights the fact that we need to develop and study degrees of contradiction and degrees of information in logical frameworks to be able to carry out correctly those reasoning tasks.

We cover in the next section some preliminary definitions for notation, and then in the following section we discuss some key dimensions for measuring inconsistent information. In the subsequent five sections, we consider five key approaches to measuring inconsistent information: Consistency-based analysis that focuses on the consistent and inconsistent subsets of a knowledgebase in Section 4; Information theoretic analysis that is an adaptation of Shannon's information measure in Section 5; Probabilistic semantic analysis that assumes a probability distribution over a set of formulae in Section 6; Epistemic actions analysis that measures the degree of information in a knowledgebase in terms of the number of actions required to identify the truth value of each atomic proposition and the degree of contradiction in a knowledgebase in terms of the number of actions needed to render the knowledgebase consistent in Section 7; and in Section 8 model-theoretic analyses that are based on evaluating a knowledgebase in terms of three or four valued models that permit an "inconsistent" truth value. We follow this range of approaches with a section covering two potential applications areas, namely multi-agent negotiation and analysis of heterogeneous sources of information. Finally, we discuss what has been achieved so far in this subject, and some possible research issues.

2 Preliminaries

For a set X, let $\wp(X)$ be the power set of X. Let $\mathcal{L}_{PS_i}$ be a language composed from a set of atoms PS and a set of logical connectives and let $\vdash_i \subseteq \wp(\mathcal{L}_{PS_i}) \times \mathcal{L}_{PS_i}$ denote the consequence relation for that language. Let $\Delta \subseteq \mathcal{L}_{PS_i}$ be a knowledgebase and let $\alpha \in \mathcal{L}_{PS_i}$ be a formula. Let $\models_i$ be a satisfaction relation for $\mathcal{L}_{PS_i}$, let $\mathsf{Models}_i(\Delta) = \{M \mid M \models_i \alpha \text{ for all } \alpha \in \Delta\}$ be the set of models for Δ in some logic i and let $\mathcal{W}_i$ be the set of models for the language $\mathcal{L}_{PS_i}$. Let $\mathsf{Consequences}_i(\Delta) = \{\alpha \mid \Delta \vdash_i \alpha\}$.

For classical logic, we drop the subscript. So $\vdash$ is the classical consequence relation and $\mathcal{L}_{PS}$ is the usual set of classical formulae formed from a set of atoms and the usual logical connectives using the usual inductive definition. If $\mathcal{L}_{PS}$ is a set of first-order formulae, then each variable in each formula is in the scope of a universal or existential quantifier as usual. For $\Delta \subseteq \mathcal{L}_{PS}$, $\mathsf{Consequences}(\Delta) = \{\alpha \mid \Delta \vdash \alpha\}$.

When it is not ambiguous we will not write the subscript PS, so we will simply write $\mathcal{L}_i$ for $\mathcal{L}_{PSi}$, and $\mathcal{L}$ for $\mathcal{L}_{PS}$.

If $\Gamma \in \wp(\mathcal{L})$, then $\mathsf{Atoms}(\Gamma)$ returns the set of atom symbols used in Γ.

Definition 1. *Let Δ be a knowledgebase and let $\vdash$ be the classical consequence relation.*

$$\mathsf{CON}(\Delta) = \{\Pi \subseteq \Delta | \Pi \nvdash \bot\}$$
$$\mathsf{INC}(\Delta) = \{\Pi \subseteq \Delta | \Pi \vdash \bot\}$$
$$\mathsf{MC}(\Delta) = \{\Pi \in \mathsf{CON}(\Delta) | \forall \Phi \in \mathsf{CON}(\Delta) \Pi \not\subset \Phi\}$$
$$\mathsf{MI}(\Delta) = \{\Pi \in \mathsf{INC}(\Delta) | \forall \Phi \in \mathsf{INC}(\Delta) \Phi \not\subset \Pi\}$$
$$\mathsf{FREE}(\Delta) = \bigcap \mathsf{MC}(\Delta)$$

Hence $\mathsf{MC}(\Delta)$ is the set of maximally consistent subsets of Δ; $\mathsf{MI}(\Delta)$ is the set of minimally inconsistent subsets of Δ; and $\mathsf{FREE}(\Delta)$ is the set of information that all maximally consistent subsets of Δ have in common. We also have the following relationship.

$$\mathsf{FREE}(\Delta) = \bigcap \mathsf{MC}(\Delta) = \Delta - \bigcup \mathsf{MI}(\Delta)$$

Example 1. Let $\Delta = \{\alpha, \neg\alpha, \alpha \to \beta, \neg\alpha \to \beta, \gamma\}$. So $\mathsf{MC}(\Delta) = \{\Phi_1, \Phi_2\}$, where $\Phi_1 = \{\alpha, \alpha \to \beta, \neg\alpha \to \beta, \gamma\}$, and $\Phi_2 = \{\neg\alpha, \alpha \to \beta, \neg\alpha \to \beta, \gamma\}$. From this, $\mathsf{FREE}(\Delta)$=$\bigcap \mathsf{MC}(\Delta)$=$\{\alpha \to \beta, \neg\alpha \to \beta, \gamma\}$, and $\mathsf{MI}(\Delta) = \{\Psi\}$, where $\Psi = \{\alpha, \neg\alpha\}$.

We can consider a maximally consistent subset of a database as capturing a "plausible" or "coherent" view on the database. For this reason, the set $\mathsf{MC}(\Delta)$ is important in many of the definitions presented in Section 4. Furthermore, we consider $\mathsf{FREE}(\Delta)$, which is equal to $\bigcap \mathsf{MC}(\Delta)$, as capturing all the "uncontroversial" information in Δ. In contrast, we consider the set $\bigcup \mathsf{MI}(\Delta)$ as capturing all the "problematical" data in Δ.

3 Dimensions of Measuring Inconsistency

To move beyond classifying a set of formulae using a binary classification (of consistent or inconsistent), we need to consider some of the dimensions we have available for measuring inconsistency.

First, there are many ways of defining inconsistency. It is a logical concept. But, there are different ways that we can view it in a language and the reasoning with that language. Inconsistency can also be viewed in the semantics. We start by considering five ways of describing inconsistency that all apply to classical logic. In classical logic, all these definitions of inconsistency coincide (i.e. when $\vdash_i$ is the classical consequence relation and $\mathsf{Models}_i(\Delta)$ is the set of classical models of Δ).

Inconsistency as Explosive Reasoning. Explosive reasoning is reasoning that allows the derivation of every formula of the language in case of inconsistency. In other words, if $\mathsf{Consequences}_i(\Delta) = \mathcal{L}_i$, then Δ is inconsistent.

Inconsistency as Conflicting Inferences. The knowledgebase Δ is inconsistent when there is the inference of both $\Delta \vdash_i \alpha$ and $\Delta \vdash_i \neg\alpha$ for some $\alpha \in \mathcal{L}_i$.

Inconsistency as Inference of a Contradiction Formulae. If the contradiction formula, denoted $\perp$, is an atom in $\mathcal{L}_i$, it can be treated in the proof theory $\vdash_i$ as logically equivalent to any inconsistent formula. So if $\Delta \vdash_i \perp$, then Δ is inconsistent. In classical logic any inconsistent formula is equivalent to any other inconsistent formula. So in an infinite classical logic language, there is an infinite number of inconsistent formulae.

Inconsistency as Trivial Reasoning. A trivial inference is an inference α from a knowledgebase Δ such that α is not a tautology and $\mathsf{Atoms}(\Delta) \cap \mathsf{Atoms}(\{\alpha\}) = \emptyset$. So if $\Delta \vdash_i \alpha$ and α is a trivial inference then Δ is inconsistent by trivial reasoning from $\vdash_i$.

Inconsistency as a Lack of a Model. If $\mathsf{Models}_i(\Delta) = \emptyset$, then Δ is inconsistent. The motivation for this is that a model is a possible coherent view of the world involving Δ. So if there is no such view, then Δ is regarded as inconsistent. This definition holds for numerous logics including classical logic.

The first description of inconsistency, i.e. "inconsistency as explosive reasoning", is a stronger definition than any of "inconsistency as inference of a contradiction formulae", "inconsistency as conflicting inferences", or "inconsistency as trivial reasoning", in the sense that an inconsistency by the first definition, is an inconsistency by the other three. Whilst the above five descriptions apply to classical logic, there are many other logics for which one or more of the above descriptions apply. Below we consider two further descriptions that apply to some logics, though neither apply to classical logic.

Inconsistency as an Inconsistent Truth Value. Let B be an inconsistent truth value. Let $\alpha \in \Delta$. If for all models of Δ, α is assigned B, then α is inconsistent, and hence Δ is inconsistent. Whilst this does not hold for classical logic, many-valued logics, and hence many-valued models, can be used to evaluate a set of classical formulae (see for example [6, 29]).

Inconsistency as Delineated Falsity. Instead of a single falsity symbol, we can adopt numerous falsity sumbols of the form $\perp_k$ and defined as $\alpha_k \rightarrow \perp_k$ for some α_k (for a brief review see [10]). This notion of inconsistency does not have the same status as the ones above. It introduces several levels of inconsistency, whereas all the other definitions above only give a dichotomy inconsistent/consistent. Note that those different levels of inconsistency can be related to the ones obtained in possibilistic logic, where the formulae deduced at a level above the inconsistency level are still safe consequences of the base, despite the presence of an inconsistency [19, 20, 8].

When we have more complex information as input, we can state several other candidate definitions for inconsistency. This extra information may be a set of plans, constraints, norms, properties, etc. Inconsistency can then be viewed operationally. Some kinds of operational definitions include:

Inconsistency as Unrealisability. If Δ is a plan or specification for something, and it is unrealisable, then Δ is inconsistent (perhaps in the context of the environment for the plan or specification);

Inconsistency as Rule Violation. If some rule is violated, then the agent, process, entity, etc. that caused the violation is inconsistent.

Inconsistency as Violation of Normality. If in a set, most of the elements have some property X, then the elements of the set that do not have this property, are inconsistent with respect to X.

It is interesting to note that these last three types of inconsistency, defined in terms of two distinct types of information — a knowledgebase plus constraints/plans/norms/properties — can be either more or less demanding than by the "classical" one. If the constraints/plans/norms/properties give some domain of interest, then the base will be considered inconsistent only if there is a conflict on the domain. In other words, we can have a conflict on variables/formulae outside of the domain without the conflict being considered inconsistent by these definitions. In this case it is less demanding than the classical definition. Conversely when the constraints/plans/norms/properties give what can be regarded as a situation of "unrealisability": For example, if the three atoms a, b and c cannot all be true at the same time, and so the base {a,b,c} is classically consistent but it is not consistent for "unrealisability".

All these ten different definitions for inconsistency offer different features of a logic that can be analysed. In this review, we can see that not all these possibilities have been considered yet.

Having selected a definition for inconsistency, together with a language and an underlying logic, there are a number of dimensions that we may wish to consider in a framework for analysing inconsistent knowledgebases in that logic. We consider some of these dimensions below.

Atomic Inconsistency. To be able to measure inconsistency, we need a formalisation of an atomic inconsistency: An indivisable and discrete representation of contradictory information. There are a number of choices depending on whether we want to take a semantic or syntactic approach, and on which underlying logic we use. The main options are to put the atomicity either on formulae or on the propositional letters. So the options we will consider here are (1) minimal inconsistent subset of formulae and (2) a propositional letter assigned with an inconsistent truth value. Another possibility which we do not consider further in this review is regarding each delineated falsity as an atomic inconsistency.

Number of Inconsistencies. Once we have a notion of atomic inconsistency, we can count them. Increasing the number of inconsistencies in a knowledgebase may or may not be a factor that increases the measure of inconsistency for that knowledgebase.

Size of Inconsistency. Once we have a notion of atomic inconsistency, we can consider the size of each atomic inconsistency, since they are not necessarily the same size. Suppose we use "minimal inconsistent subset" as the definition for an atomic inconsistency. Suppose also that Δ_1 and Δ_2 are minimal inconsistent subsets of some knowledgebase Δ, and $|\Delta_1| \leq |\Delta_2|$ holds, then Δ_2 is a bigger inconsistency than Δ_1. This is only one way we may choose to evaluate the size of an inconsistency. Increasing the size of inconsistency may or may not be a factor that increase the measure of inconsistency.

Degree of Information. Measuring the amount of information in a message or source is well established with proposals such as Shannon's information theory. In the usual applications of Shannon's information theory, inconsistent information contains no information. This coincides with a classical logic perspective of inconsistency (i.e. there are no models of inconsistent information, and as a result it represents no information). However, information about the real-world frequently, or normally, incorporates inconsistency, and yet it is still informative. So the intuition that inconsistent information contains useful information, leads to proposals for measuring the degree of information in the context of inconsistency.

Further dimensions that we may consider include the following two. The first of these could be described as a composite measure, using both degree of information and degree of contradiction, and the second of these requires further (meta-level) information.

Ratio of Information to Noise. When considering inconsistent information, if there is a relatively large amount of information when compared with the amount of inconsistency, then that source is likely to be more acceptable than a source that has a relatively low amount of information when compared with the amount of inconsistency.

Significance of Inconsistency. As an illustration of the need to evaluate significance, consider two news reports on a World Cup match, where the first report says that Brazil beat Germany 2-0, and the second report says that Germany beat Brazil 2-0. This is clearly a significant inconsistency. Now consider two news reports on the same football match, where the first report says that the referee was Pierluigi Collina and the second report says that the referee was Ubaldo Aquino. This inconsistency would normally be regarded as relatively insignificant.

Amongst the five approaches to measuring inconsistent information in this review, namely consistency-based analysis, information-theoretic analysis, analysis of probabilistic semantic, analysis of epistemic actions, and model-theoretic analysis, we see these dimensions drawn out.

4 Consistency-Based Analyses

One of the most obvious strategies for handling inconsistency in a knowledgebase is to reason with consistent subsets of the knowledgebase. This is closely related to the approach of removing information from the knowledgebase that is causing an inconsistency (see for example [52, 7, 21]).

To measure the information, and the degree of inconsistency, we can take cardinality of the Δ and $\mathsf{MI}(\Delta)$ sets as the basis of an analysis. We can use this for the following ratio that captures the relative incompatibility of the formulae in the knowledgebase.

Definition 2. *The* **incompatibility ratio** *for a knowledgebase* $\Delta \subseteq \mathcal{L}_{PS}$ *is defined as follows.*

$$\frac{|\mathsf{MI}(\Delta)|}{|\Delta|}$$

Example 2. Let $\Delta = \{\alpha, \neg\alpha, \beta, \neg\beta, \gamma, \delta, \gamma \wedge \delta\}$

$$\frac{|\mathsf{MI}(\Delta)|}{|\Delta|} = \frac{2}{7}$$

Whilst this ratio provides an abstraction of the conflicts in the information in Δ, it says nothing about the relative size of the minimal inconsistent subsets, or the overlaps between members of $\mathsf{MI}(\Delta)$. Also the syntax sensitivity can be problematical.

Example 3. Let $\Delta_1 = \{\alpha \wedge \beta, \neg\alpha \wedge \neg\beta\}$ and $\Delta_2 = \{\alpha \wedge \neg\alpha, \beta \wedge \neg\beta\}$.

$$\frac{|\mathsf{MI}(\Delta_1)|}{|\Delta_1|} = \frac{1}{2} \qquad \frac{|\mathsf{MI}(\Delta_2)|}{|\Delta_2|} = \frac{2}{2}$$

These shortcomings in part stem from this measure being insufficiently fine grained. To address this, we will now review an approach based on scoring functions that provides a deeper consistency-based analysis of the inconsistencies arising in a set of formulae.

For a knowledgebase Δ, a scoring function S is from $\wp(\Delta)$ into the natural numbers defined so that $S(\Gamma)$ gives the number of minimally inconsistent subsets of Δ that would be eliminated if the subset Γ was removed from Δ [34]. This characterization offers an alternative means for articulating, in general terms, the nature of inconsistency in a set of formulae. Knowledgebases can be compared using their scoring functions giving an ordering relation over databases that can be described as “more conflicting than”.

Definition 3. *Let* $\Delta \subseteq \mathcal{L}_{PS}$. *Let* S *be the* **scoring function** *for* Δ *defined as follows, where* $S : \wp(\Delta) \mapsto \mathbb{N}$ *and* $\Gamma \in \wp(\Delta)$

$$S(\Gamma) = |\mathsf{MI}(\Delta)| - |\mathsf{MI}(\Delta - \Gamma)|$$

The scoring function for a database is an abstraction of the information we have about the database, and it says much about the inconsistencies arising in the database.

Example 4. Let $\Delta = \{\alpha, \neg\alpha, \beta \wedge \neg\beta\}$, where S is the scoring function for Δ, defined as follows:

$$\begin{array}{lll} S(\{\alpha\}) = 1 & S(\{\neg\alpha\}) = 1 & S(\{\beta \wedge \neg\beta\}) = 1 \\ S(\{\alpha, \neg\alpha\}) = 1 & S(\{\alpha, \beta \wedge \neg\beta\}) = 2 & S(\{\neg\alpha, \beta \wedge \neg\beta\}) = 2 \\ & S(\{\alpha, \neg\alpha, \beta \wedge \neg\beta\}) = 2 & \end{array}$$

Example 5. Let $\Delta = \{\alpha \wedge \neg\alpha, \beta, \gamma\}$, where S is the scoring function for Δ, defined as follows:

$$\begin{array}{lll} S(\{\alpha \wedge \neg\alpha\}) = 1 & S(\{\beta\}) = 0 & S(\{\gamma\}) = 0 \\ S(\{\alpha \wedge \neg\alpha, \beta\}) = 1 & S(\{\alpha \wedge \neg\alpha, \gamma\}) = 1 & S(\{\beta, \gamma\}) = 0 \\ & S(\{\alpha \wedge \neg\alpha, \beta, \gamma\}) = 1 & \end{array}$$

We can make a few simple observations regarding scoring functions. Where S is the scoring function for Δ, $S(\cup\mathsf{MI}(\Delta)) = S(\Delta) = |\mathsf{MI}(\Delta)|$ and $S(\mathsf{FREE}(\Delta)) = 0$. Also from the scoring function for a database Δ, it is straightforward to calculate the cardinality of $\mathsf{FREE}(\Delta)$ and $\cup\mathsf{MI}(\Delta)$. However, there is no simple way for determining the cardinality of the set of maximally consistent subsets of a database directly from the scoring function for the database.

Proposition 1. *Let $\leq$ be the usual ordering relation over $\mathbb{N}$. For $\Gamma_i, \Gamma_j \in \wp(\Delta)$, where S is the scoring function for Δ,*

$$S(\Gamma_i \cap \Gamma_j) \leq \mathsf{min}(\{S(\Gamma_i), S(\Gamma_j)\})$$

$$\mathsf{max}(\{S(\Gamma_i), S(\Gamma_j)\}) \leq S(\Gamma_i \cup \Gamma_j)$$

Note, $S(\Gamma_i) + S(\Gamma_j) \leq S(\Gamma_i \cup \Gamma_j)$ does not necessarily hold as illustrated below.

Example 6. Let S be the scoring function for Δ, and let $\Gamma_1 = \{\neg\alpha, \alpha \wedge \beta\}$, and let $\Gamma_2 = \{\beta, \alpha \wedge \neg\beta\}$, and let $\Delta = \Gamma_1 \cup \Gamma_2$. So $S(\Gamma_1) = S(\Gamma_2) = 3$, but $S(\Gamma_1 \cup \Gamma_2) = 4$.

We can compare databases using the scoring function for each database. For this we define score orderings.

Definition 4. *A* **score ordering***, denoted $\leq$, is defined as follows*[1]*. Assume Δ_i and Δ_j are of the same cardinality and S_i is the scoring function for Δ_i,*

[1] Note, we are now using the $\leq$ symbol for the usual ordering over the natural numbers and as defined here for an ordering over score functions. Hopefully, this overloading of the symbol will not cause confusion.

and S_j is the scoring function for Δ_j. $S_i \leq S_j$ holds iff there is a bijection $f : \wp(\Delta_i) \mapsto \wp(\Delta_j)$ such that the following condition is satisfied:

$$\forall \Gamma \in \wp(\Delta_i), S_i(\Gamma) \leq S_j(f(\Gamma))$$

Note, $S_i < S_j$ iff $S_i \leq S_j$ and $S_j \not\leq S_i$. Also, $S_i \simeq S_j$ iff $S_i \leq S_j$ and $S_j \leq S_i$. We say Δ_j is **more inconsistent** *than Δ_i iff $S_i \leq S_j$.*

Example 7. Let $\Delta_1 = \{\alpha, \neg\alpha\}$ and $\Delta_2 = \{\alpha, \beta \wedge \neg\beta\}$. Let S_1 be the scoring function for Δ_1 and S_2 be the scoring function for Δ_2, and so $S_2 < S_1$.

$$\begin{array}{ll} S_1(\{\alpha\}) = 1 & S_2(\{\alpha\}) = 0 \\ S_1(\{\neg\alpha\}) = 1 & S_2(\{\beta \wedge \neg\beta\}) = 1 \\ S_1(\{\alpha, \neg\alpha\}) = 1 & S_2(\{\alpha, \beta \wedge \neg\beta\}) = 1 \end{array}$$

Example 8. Consider $\Delta_1 = \{\alpha \wedge \neg\alpha, \beta, \gamma\}$ and $\Delta_2 = \{\alpha \wedge \neg\alpha, \beta \wedge \neg\beta, \delta\}$. If S_1 is the scoring function for Δ_1, and S_2 is the scoring function for Δ_2, then $S_1 < S_2$.

We can consider scoring functions as giving information about the overlaps of the minimally inconsistent subsets. For example, for Δ_i and Δ_j, if $|\Delta_i| = |\Delta_j|$ and $|\mathsf{MI}(\Delta_i)| = |\mathsf{MI}(\Delta_j)|$ and $S_i \leq S_j$ then the inconsistencies are more overlapping in Δ_j. In other words, more of the formulae are in more minimally inconsistent subsets. In case we want to compare sets of different cardinality, we can add dummy propositions to the smaller set to make it the same size as the larger set. These dummy propositions are literals that do not appear elsewhere and so can be assumed to not be in any of the minimally inconsistent subsets of the database.

For each $n \in \mathbb{N}$, the score ordering $\leq$ over knowledgebases is reflexive and transitive, but not antisymmetric. The following result shows in part how a score ordering can be viewed as an aggregation of parameters including the relative number of minimally inconsistent formulae and the relative number of free formulae.

Proposition 2. *If $|\Delta_i| = |\Delta_j|$, and S_i is the scoring function for Δ_i, and S_j is the scoring function for Δ_j, then*

$$S_i \leq S_j \textit{ implies } |\mathsf{MI}(\Delta_i)| \leq |\mathsf{MI}(\Delta_j)|$$
$$S_i \leq S_j \textit{ implies } |\mathsf{FREE}(\Delta_i)| \geq |\mathsf{FREE}(\Delta_j)|$$

Note, the converse does not hold.

With the same assumptions as those for Proposition 2, we do not get that $S_i \leq S_j$ implies $|\mathsf{MC}(\Delta_i)| \leq |\mathsf{MC}(\Delta_j)|$ or that it implies $|\mathsf{MC}(\Delta_i)| \geq |\mathsf{MC}(\Delta_j)|$. This is captured in the following example.

Example 9. Consider $\Delta_1 = \{\alpha, \beta\}$ and $\Delta_2 = \{\alpha, \neg\alpha\}$. So $S_1 \leq S_2$ and $|\mathsf{MC}(\Delta_1)| \leq |\mathsf{MC}(\Delta_2)|$. Now consider $\Delta_3 = \{\alpha, \neg\alpha\}$ and $\Delta_4 = \{\beta \wedge \neg\beta, \gamma \wedge \neg\gamma\}$. So $S_3 \leq S_4$ and $|\mathsf{MC}(\Delta_3)| \geq |\mathsf{MC}(\Delta_4)|$.

Clearly, scoring functions are syntax sensitive in the sense that we may have two knowledgebase Δ_1 and Δ_2 where $\mathsf{Consequences}(\Delta_1) = \mathsf{Consequences}(\Delta_2)$ and S_1 is the scoring function for Δ_1 and S_2 is the scoring function for Δ_2, but $S_1(\Delta_1) \neq S_2(\Delta_2)$. Scoring functions may also be regarded as being prone to semantic insensitivity. To illustrate semantic insensitivity, consider the following two examples.

Example 10. Consider Δ_1 and Δ_2 below. Let S_1 be the scoring function for Δ_1 and S_2 be the scoring function for Δ_2.

$$\Delta_1 = \{\alpha, \neg\alpha\}$$
$$\Delta_2 = \{\alpha \wedge \beta, \neg\alpha \wedge \beta\}$$

Here, $S_1 \simeq S_2$ and so the scoring functions do not differentiate Δ_1 and Δ_2. Yet it could be argued that semantically Δ_2 implies more (such as if paraconsistent logic inference were used) than Δ_1.

Example 11. Consider Δ_1 and Δ_2 below. Let S_1 be the scoring function for Δ_1 and S_2 be the scoring function for Δ_2.

$$\Delta_1 = \{\alpha \wedge \beta \wedge \gamma, \alpha \wedge \neg\beta \wedge \gamma, \neg\alpha \wedge \beta \wedge \neg\gamma\}$$
$$\Delta_2 = \{\alpha \wedge \beta \wedge \gamma, \alpha \wedge \neg\beta \wedge \gamma, \neg\alpha \wedge \beta \wedge \gamma\}$$

Here, the formulae in Δ_1 and Δ_2 are pairwise inconsistent, and the resulting scoring functions are such that $S_1 \simeq S_2$. It may be argued that Δ_2 is less inconsistent than Δ_1 since all formulae in Δ_2 agree on γ.

In response to the arguments raised in Example 10 and 11, we believe that this kind of semantic insensitivity is useful in some applications. We believe that when a connective is used, it is used with some intent. So for example, whilst $\alpha \wedge \beta$ and α, β are semantically equivalent, we may need to differentiate them also. This intent depends on the applications area, but to illustrate in negotiation, consider a strategy for weakening the preferences (represented by a set of classical formulae) of an agent is take a subset of the preferences. So if an agent starts with $\{\alpha \wedge \beta\}$ as its preferences then the only possible weakening (using the $\subseteq$ relation) is $\{\}$. Whereas if the agent starts with $\{\alpha, \beta\}$ then weakenings also include $\{\alpha\}$ and $\{\beta\}$. In this application, the preference $\alpha \wedge \beta$ is intended to mean that $\alpha \wedge \beta$ must occur together, and so if the preference α is dropped then so is the preference β.

In fact, this question is related to the status of the "comma" connective. In classical logic (in the consistent case) $\{\alpha \wedge \beta\}$ and $\{\alpha, \beta\}$ are logically equivalent. That shows that the comma in the second knowledgebase has exactly the same meaning as a conjunction. But as soon as the knowledgebase is not consistent, a lot of approaches give a different meaning to the comma and to the conjunction. As explained above, this can be sensible if the $\wedge$ connective means that the conjuncts must absolutly occur together, whereas it is not the case with the comma. This difference can be very intuitive, but it is not mandatory. And this choice leads to different approaches.

The general conclusion we draw from this discussion is that the syntax sensitivity, and the semantic insensitivity, found in scoring functions is useful in some applications.

5 Information Theoretic Analyses

First we consider how information theory can be used to measure the information content of propositional formulae [70].

Table 1. A 3×3 grid denoting 9 possible locations for an object

	β_1	β_2	β_3
α_1	$\times$		$\times$
α_2			
α_3			

Example 12. Let ϕ be a formula in the classical language composed of the following propositional letters $\{\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3\}$. Now consider that ϕ represents the location of an object where there are 9 possible positions in a 3×3 grid (cf. Table 1). Information can be collected on the position in the grid. Now if we receive two messages: The first states that the position is α_1 and the second states that the position is $\neg\beta_2$. From these two statements, we can conclude that the position is $\alpha_1 \wedge (\beta_1 \vee \beta_3)$. This is represented by the $\times$ symbol in Table 1.

The basic idea behind Shannon's measure of information is that information eliminates possibilities. The more unlikely a piece of information, the more information is conveyed when that piece of information is asserted.

Definition 5. *Let ϕ be a piece of information, and let $P(\phi)$ be the probability of ϕ occuring.* **Shannon's information measure** I *is*

$$I(\phi) = -logP(\phi)$$

We can illustrate the use of Shannon's Information measure by the following example.

Example 13. Returning to Example 12, we can use the 3×3 grid as a probability space and we can assume a uniform distribution over this space (i.e. each position in the grid is equally probable). Using Definition 5 for $\phi = \alpha_1$ and $\phi' = \alpha_1 \wedge \neg\beta_2$, we get

$$I(\phi) = -log\frac{3}{9} = 0.48 \qquad I(\phi') = -log\frac{2}{9} = 0.65$$

Information theory can be used to measure the information content of sets of consistent formulae. The information in a set Γ, composed from n different atom symbols, is the logarithm of the number of models (2^n) divided by the number of models of Γ [48]. This idea can be traced back to Kemeny [35] and Hintikka [27], so we will call this measure the Kemeny and Hintikka measure of information.

Definition 6. *Let Γ be a consistent set of formulae, let n be the number of atoms in the language $\mathcal{L}_{PSi}$, and let $\mathsf{Models}(\Gamma)$ denote the collection of models for Γ. The information value of Γ is defined by the following equation.*

$$I(\Gamma) = log_2 \frac{2^n}{|\mathsf{Models}(\Gamma)|}$$

Rewriting this equation, we get the following.

$$I(\Gamma) = n - log_2 |\mathsf{Models}(\Gamma)|$$

Notice that with this definition, if the set of formulae Γ is inconsistent, then the measure of information does not work. To address this, Lozinskii extends this approach to measure the information content of sets of inconsistent formulae. The information in a set Γ, composed from n different atom symbols, is the logarithm of the number of models (2^n) divided by the number of models for the maximum consistent subsets of Γ.

Definition 7. *Let Γ be a consistent set of formulae, let n be the number of atoms in the language $\mathcal{L}_{PSi}$, and let $\mathsf{MC}(\Gamma)$ be the set of maximally consistent subsets of Γ (see Definition 1). For each $\Delta \in \mathsf{MC}(\Gamma)$, if $M(\Delta)$ is the collection of models of Δ, then the collection of quasi-models is defined by*

$$U(\Gamma) = \bigcup \{\mathsf{Models}(\Delta) \mid \Delta \in \mathsf{MC}(\Gamma)\}$$

The information value of Γ is defined by the following equation

$$I_l(\Gamma) = n - log_2 |U(\Gamma)|$$

This measure increases with additions of consistent information and decreases with additions of inconsistent information.

Example 14. For $\Delta = \{\alpha \vee \beta, \alpha \vee \neg\beta, \neg\alpha \wedge \gamma\}$, $\Gamma = \Delta \cup \{\neg\gamma\}$, and $\Gamma' = \Delta \cup \{\delta\}$

$$I_l(\Gamma) < I_l(\Delta) \qquad\qquad I_l(\Gamma') > I_l(\Delta)$$

However, the approach is syntax sensitive, in the sense discussed in Section 4, as illustrated by the following example.

Example 15. For $\Delta = \{\alpha \vee \beta, \alpha \vee \neg\beta, \neg\alpha \wedge \gamma\}$, and $\Delta' = \{\alpha \vee \beta, \alpha \vee \neg\beta, \neg\alpha, \gamma\}$, but $I_l(\Delta') < I_l(\Delta)$.

To address this syntax sensitivity, a normal form can be used for application of Lozinskii's measure. One proposal is to rewrite all formulae into conjunctive normal form, and then exhaustively apply conjunction elimination and resolution [70]. Returning to Example 15, if we use this normal form of the knowledgebases, then they have the same result using Lozinskii's measure.

It should be noted, whether or not we use a normal form, this information-theoretic approach does not provide a direct evaluation of inconsistency since for example, the value for $\{\alpha\}$ is the same as for $\{\alpha, \neg\alpha, \beta\}$. As a result, we stress this approach provides a measure of information that may be inconsistent rather than a measure of the inconsistencies in the information.

An interesting application of Lozinskii's measure is in a form of belief revision [50].

6 Analysis of Probabilistic Semantics

There is an inconsistency analysis framework, with a probabilistic semantics, that assigns a measure of consistency in the range $[0, 1]$ for each set of propositional formulae [37, 39]. For a set of formulae that contains a formula that is contradictory (logically equivalent to falsity), the measure of consistency is 0. For a set of formulae that is consistent, the measure of consistency is 1.

For any set of formulae, the measure of consistency is directly proportional to the size of the minimal inconsistent subsets. This conceptualizes the intuition that the more formulae required to obtain an inconsistency, the more tolerable the set becomes. Since, any formula equivalent to falsity causes that set to have a measure of consistency of 0, the framework collapses a number of interesting, different kinds of knowledgebase, to the same value. In this sense, the framework is less-fined grained than the other measures presented in the next sections in the propositional case.

On the other hand this measure is the only one among those presented here that takes into account the number of formulae required to lead to the inconsistency.

It is also possible to define two information measures based on the measure of consistency [39].

6.1 Probabilistic Measure of Consistency

First one needs to define a probability function over formulae:

Definition 8. *A* probability function *on $\mathcal{L}$ is a function $P : \mathcal{L} \rightarrow [0, 1]$ s.t. :*

- *if $\models \alpha$, then $P(\alpha) = 1$*
- *if $\models \neg(\alpha \wedge \beta)$, then $P(\alpha \vee \beta) = P(\alpha) + P(\beta)$*

See [55] for more details on this definition. In the finite case, this definition gives a probability distribution on interpertations, and the probablity of a formula is the sum of the probabilities of its models[2].

Then the measure of consistency is defined as [37] :

Definition 9. *Let Δ be a knowledgebase.*

- *Δ is $\eta-$consistent ($0 \leq \eta \leq 1$) if there exists a probability function P such that $P(\alpha) \geq \eta$ for all $\alpha \in \Delta$.*
- *Δ is maximally $\eta-$consistent if η is maximal (i.e. if $\gamma > \eta$ then Δ is not $\gamma-$consistent).*

So the notion of maximally $\eta-$consistency can be used as a measure of contradiction. This is a direct formalisation of the fact that the more formulae are

[2] It can be also defined, like in[37, 39], in terms of the logical formulae corresponding to the models of the knowledgebase (maximally coherent conjunction of literals of $\mathcal{L}$).

required to produced the inconsistency, the less the inconsistency is important. As easily seen, in a finite setting, a knowledgebase Δ is $0-$consistent if and only if there is a contradictory formula (i.e. a formula logically equivalent to falsity) in it. And Δ is $1-$consistent if and only if the knowledgebase is consistent. Let us see some examples to illustrate the non-extremal cases:

Example 16. $\Delta = \{a, b, \neg a \vee \neg b\}$ is maximally $\frac{2}{3}-$consistent. $\Delta = \{a \wedge b, \neg a \wedge \neg b, a \wedge \neg b\}$ is maximally $\frac{1}{3}-$consistent, whereas each of its subsets of cardinality 2 is maximally $\frac{1}{2}-$consistent.

For minimal inconsistent sets of formulae, computing the probabilistic measure of consistency is easy.

Proposition 3. *If $\Gamma \in \mathsf{MI}(\Delta)$, then Γ is maximally $\frac{|\Gamma|-1}{|\Gamma|}-$consistent.*

For a general knowledgebase, there is no direct way to compute it. But a lower bound can be stated:

Proposition 4. *If Δ is finite and $\Gamma \subseteq \Delta$ is a smallest minimally inconsistent subset of Δ, then Δ is $\frac{|\Gamma|-1}{|\Delta|}-$consistent.*

In fact, as underlined in [37], it can be computed using the simplex method.

6.2 Probabilistic Measures of Information

From this probabilistic measure of consistency, one can define two probabilistic measures of information [39]. Let us first give some definitions:

Definition 10. *Let Δ be a knowledgebase.*

- *A probability function P is* Pareto optimal *for Δ if there is no probability function P^* such that $P^*(\alpha) \geq P(\alpha)$ for all $\alpha \in \Delta$ and $P^*(\beta) > P(\beta)$ for one $\beta \in \Delta$.*
- *A probability function P is* $\Delta-$consistent *if Δ is maximally $\eta-$consistent and $P(\alpha) \geq \eta$ for all $\alpha \in \Delta$.*
- *A probability function P is* Rawls optimal *for Δ if it $\Delta-$consistent and Pareto optimal for Δ.*

Definition 11. *Let us note respectively R_Δ and V_Δ, the set of Rawls optimal probability functions for Δ and the set of $\Delta-$consistent probability functions.*

Proposition 5. *Let Δ be a knowledge base. A probability function is Pareto/Rawls optimal only if $P(\alpha) = 1$ for all $\alpha \in \mathsf{FREE}(\Delta)$.*

The entropy [67] of a probability function P is defined as

$$H(P) = -\sum_{\omega \in \mathcal{W}} P(\omega) \log_2 P(\omega)$$

Let X be a set of probabiliy functions, then $ME(X)$ is a random maximum entropy function from X. Then the two probabilistic measures of information are:

Definition 12. *Let Δ be a knowledgebase.*

- $I_{k1}^{\mathcal{L}}(\Delta) = |\mathcal{L}| - H(ME(R_\Delta))$
- $I_{k2}^{\mathcal{L}}(\Delta) = |\mathcal{L}| - H(ME(V_\Delta))$

These two definitions are two direct generalization of Kemeny and Hintikka measure of information (see definition 6) that do not trivialize to 0 when the knowledgebase is not consistent.

In terms of this probabilistic semantics the information measure of Kemeny and Hintikka of a knowledgebase is the size of the language minus the entropy of the probabilistic function that has a maximum entropy while giving a probability 1 to all of the formulae of the knowledgebase.

This view trivializes when the knowledgebase is not consistent, since, in this case, it is not possible to give a probability 1 to all the formulae of the knowledgebase. The two definitions proposed in Definition 12 are the two more intuitive modifications of this intuition to fit the inconsistent case. The second one ($I_{k2}^{\mathcal{L}}$) takes the maximum entropy probabilistic function that maximizes the minimum probability of the formulae of the knowledgebase. The first one ($I_{k1}^{\mathcal{L}}$) takes the maximum entropy probabilistic function that gives a probability 1 to the maximum of formulae of the knowledgebase (the free formulae of the knowledgebase) and a maximum probability to each of the other formulae of the knowledgebase. So, in both cases, the requirement of giving probability 1 to all the formulae of the knowledgebase, is "minimally" changed.

Let us see what those measures give on some simple example.

Example 17. $\Delta = \{a, b, c, \neg a \wedge \neg b\}$. Δ is maximally $\frac{1}{2}-$consistent. As $\mathsf{FREE}(\Delta) = \{c\}$, for the first information measure, the probabilistic function must satisfy $P(c) = 1$. The only probability distribution on interpretations that is Rawls optimal is $P(\{\neg a, \neg b, c\}) = \frac{1}{2}$ and $P(\{a, b, c\}) = \frac{1}{2}$. So $I_{k1}^{\mathcal{L}}(\Delta) = 2$. For the second information measure, we only consider $\Delta-$consistent probabilistic functions, and the one of maximum entropy is the one that gives a probability of $\frac{1}{4}$ to all of $\{\neg a, \neg b, c\}$, $\{\neg a, \neg b, \neg c\}$, $\{a, b, c\}$ and $\{a, b, \neg c\}$. So $I_{k2}^{\mathcal{L}}(\Delta) = 1$. Note that on this example Lozinskii's measure of information gives $I_l^{\mathcal{L}}(\Delta) = 2$.

Knight advocates the superiority of those two measures on the one proposed by Lozinskii, since they take into account the knowledgebase as a whole (i.e. all the formulae of the knowledgebase), whereas Lozinskii's, which is based on maximal consistent subsets, takes into account only some subsets of the knowledgebase.

Indeed, the two approaches (Lozinskii and Knight) give significant different results on some illustrating case. See the following example.

Example 18. Let α be a non-tautological consistent formula, and let $\Delta = \{\alpha, \neg\alpha\}$.

Then $I_l^{\mathcal{L}}(\Delta) = 0$ for all α. But, this is not the case for Knight's measures of information. Since they are based on a probability distribution on interpetations and on maximum entropy, the result depends on the number of interpretations in α and its negation.

For example if α is a conjunction of atoms $\alpha = a_1 \wedge \ldots \wedge a_k$. Then for $k = 1$, then $I^{\mathcal{L}}_{k1}(\Delta) = I^{\mathcal{L}}_{k2}(\Delta) = 0$, for $k = n$, then $I^{\mathcal{L}}_{k1}(\Delta) = I^{\mathcal{L}}_{k2}(\Delta) = |\mathcal{L}| - \frac{1}{2} - \frac{1}{2}\log_2 \frac{1}{n}$. So $I^{\mathcal{L}}_{k1}$ and $I^{\mathcal{L}}_{k2}$ increase with the number of conjuncts.

To know what behaviour is the more intuitive one is not an easy task. We think that both behaviours have pro and cons.

But both approaches are highly syntax sensitive, in particular, they make a distinction between $\{a \wedge b\}$ and $\{a, b\}$. This is not the case with some of the other approaches presented in the rest of this paper. Let us see this on the following example:

Example 19. Let $\Delta = \{a, b, \neg a \vee \neg b\}$ and $\Delta' = \{a \wedge b, \neg a \vee \neg b\}$. Then $I^{\mathcal{L}}_{k1}(\Delta) = I^{\mathcal{L}}_{k2}(\Delta) = 0.42$ and $I^{\mathcal{L}}_{l}(\Delta) = 1$. While $I^{\mathcal{L}}_{k1}(\Delta') = I^{\mathcal{L}}_{k2}(\Delta') = 0.21$ and $I^{\mathcal{L}}_{l}(\Delta) = 0$.

Note that Lozinskii's approach is immune to the presence of a contradictory formula (i.e. adding a contradictory formula to a knowledgebase does not change the measure of information). $I^{\mathcal{L}}_{k1}(\Delta)$ is dependent to the addition of contradictory formulae: adding a contradictory formulae decreases the measure of information. So it seems that this measure melds together information and contradiction. Whereas $I^{\mathcal{L}}_{k2}(\Delta)$ trivializes as soon as the knowledgebase contains a contradictory formula.

Note that the three measures by Lozinskii and Knight trivialize as soon as the knowledgebase contains only one formula that is contradictory. This is not the case with the measures presented in the following sections.

7 Analysis of Epistemic Actions

An alternative approach to quantifying degrees of information and contradiction in propositional logic is based on a framework of "epistemic actions" [40]. Each epistemic action (called also *test*) reduces inconsistency and/or gains information. The degree of information in a knowledgebase is based on the number (or the cost) of actions needed to identify the truth value of each atomic proposition: The lower the cost, the more information is contained in the base. The degree of contradiction in a knowledgebase is based on the number (or the cost) of actions needed to render the knowledgebase classically consistent. Both measurements are dependent on the language, logic, and tests used.

So this framework does not define a unique measure of contradiction (and information), but a wide familly of such measures. Since instantiating the parameters of the framework allows to define different measures. The main parameter is the underlying logic. Each propositional logic that satisfies some basic requirements can be used here, leading to different definitions of measures of contradiction (and measures of information). Another important parameter is the available epistemic actions (called *tests* in the following). Specifying the set of available tests, allows us to make a distinction between atoms/formulae. It can for example be used to state that only some atoms are of interest (i.e. contradiction or lack of information on the remaining atoms is not important), or that some atoms/formulae are more important (or more difficult to test) than other ones. We will present here only the measures of contradiction, see [40] for more details and for the definition of the corresponding measures of information.

One of the aim of this approach is to be able to say something on a single contradictory formula. Another way to express this idea is to say that in this approach the connector $\wedge$ is the same as the comma connective (i.e. $\{\alpha, \beta\}$ must be considered exactly as $\{\alpha \wedge \beta\}$). So in this section we will consider that the knowledgebase is a single formula (since, with the above hypothesis, we can take equivalently the formula that is the conjunction of all formulae of the knowledgebase).

So this approach needs an underlying propositional logic $\mathcal{L}_{PSi}$ that is required to have the following components :

1. A *consequence relation* $\models_i$ on $\mathcal{L}_{PSi} \times \mathcal{L}_{PSi}$.
2. An *acceptance function* $A_i \subseteq \mathcal{L}_{PSi} \times \mathcal{L}_{PSi}$: $A_i(\Delta, \alpha)$ means that given the knowledgebase Δ, α is accepted as true information (we say that Δ *accepts* α). By default, acceptance is defined by: $A_i(\Delta, \alpha)$ iff ($\Delta \models_i \alpha$ and $\Delta \not\models_i \neg\alpha$). We say that Δ is *informative about* α iff exactly one of $A_i(\Delta, \alpha)$ or $A_i(\Delta, \neg\alpha)$ holds, and that Δ is *fully informative* iff for any $\alpha \in \mathcal{L}_{PSi}$, Δ is informative about α.
3. A *contradiction indicator* $C_i \subseteq \mathcal{L}_{PSi} \times \mathcal{L}_{PSi}$: if $C_i(\Delta, \alpha)$ holds, then we say that Δ *is contradictory about* α. By default, we define $C_i(\Delta, \alpha)$ iff ($\Delta \models_i \alpha$ and $\Delta \models_i \neg\alpha$). Δ is said to be *contradiction-free* iff for every $\alpha \in \mathcal{L}_{PSi}$, we do not have $C_i(\Delta, \alpha)$.
4. A *weak revision* operator $\star : \mathcal{L}_{PSi} \times \mathcal{L}_{PSi} \to \mathcal{L}_{PSi}$: $\Delta \star \alpha$ represents the new knowledgebase obtained once taking account of the observation α into the knowledgebase Δ. For the sake of generality, we are not very demanding about $\star$. The only requirement is that $\Delta \star \alpha \models_i \alpha$, which expresses that our tests are assumed reliable (each test outcome must be true in the actual world). In the following we will simply refer to these operators as *revision* operators (omitting the *weak*).

Let us now define the set of tests (a test is an action that allows to truthfully know the truth value of a formula) that will be allowed to be used for computing the measure of contradiction.

Definition 13. *A* test context $\mathcal{C}_{\mathcal{L}_{PSi}}$ *(w.r.t. $\mathcal{L}_{PSi}$) is a pair $\langle T, c\rangle$ where T is a finite set of tests and c is a cost function from T to $\mathbb{N}^*$ (the set of strictly positive integers). The* outcome *to any test $t_\alpha \in T$ is one of α, $\neg\alpha$, where $\alpha \in \mathcal{L}_{PSi}$. We say that t_α is the*[3] *test about α . A context is said to be :*

– standard *iff $\forall t_\alpha \in T$, we have $c(t_\alpha) = 1$ (every test has a unit cost).*
– universal *iff for every $\alpha \in \mathcal{L}_{PSi}$, there is a test $t_\alpha \in T$.*
– atomic *iff the testable formulae are exactly the atoms of the language ($t_x \in T$ iff $x \in PS$).*

[3] It is assumed, without loss of generality, that at most one test t_α of T is about α for each $\alpha \in \mathcal{L}_{PSi}$.

Definition 14. *Given a test context $\mathcal{C}_{\mathcal{L}_{PS_i}}$, a* test plan π *is a finite binary tree; each of its non-terminal nodes is labelled with a test action t_α; the left and right arcs leaving every non-terminal node labelled with t_α are respectively labelled with the outcomes α and $\neg\alpha$. An* (outcome) trajectory $\langle o_1, \ldots, o_n \rangle$ *with respect to π is the sequence of test outcomes on a branch of π. The* cost *of a trajectory $\langle o_1, \ldots, o_n \rangle$ with respect to π is defined as $\sum_{i=1}^{n} c(t_{\alpha_i})$, where each t_{α_i} is the test labelling the node of π reached by following the path $\langle o_1, \ldots, o_{i-1} \rangle$ from the root of π.*

Test plans are the basic tool used to determinate the tests needed to remove the inconsistency. It is the number (or more generally the cost) of the tests needed that will give the measure of contradiction for a knowledgebase.

Definition 15. *Let π be a test plan and Δ the initial knowledgebase.*

- *The application of π on Δ is the tree apply(π, Δ), isomorphic to π, whose nodes are labelled with knowledgebases defined inductively as follows:*
  - *the root ϵ of apply(π, Δ) is labelled with $\Delta(\epsilon) = \Delta$;*
  - *let n be a node of apply(π, Δ), labelled with the knowledgebase $\Delta(n)$, whose corresponding node in π is non-terminal and labelled with t_α; then n has two children in apply(π, Δ), labelled respectively with $\Delta(n) \star \alpha$ and $\Delta(n) \star \neg\alpha$.*
- π purifies α given Δ *iff for every terminal node n of apply(π, Δ), $\Delta(n)$ is not contradictory about α (i.e., not $C_L(\Delta(n), \alpha)$).*
- π (fully) purifies Δ *iff it eliminates all contradictions in Δ, i.e., iff for any terminal node n of apply(π, Δ), $\Delta(n)$ is contradiction-free.*

The degree of contradiction of Δ measures the minimal effort necessary to purify Δ.

Note that, clearly enough, it can be the case that there is no plan to purify a formula (if the test context is not atomic and not universal).

Definition 16. *Let us define the* cost $c(\pi)$ *of a test plan π as the maximum of the costs of its trajectories. Then the* degree of contradiction *of Δ is defined by $d_C(\Delta) = \mathsf{min}(\{c(\pi) \mid \pi \text{ purifies } \Delta\})$. When no plan purifies Δ, we let $d_C(\Delta) = +\infty$.*

In the previous definition, we actually define *pessimistic degrees* of contradiction (because the cost of a plan is defined as the *maximum* cost among its trajectories); this principle, consisting in assuming the worst outcome, is known in decision theory as *Wald criterion*. Other criteria could be used instead, such as the optimistic criterion obtained by replacing *max* by *min*. Also interesting, the criterion obtained by first using *max* and then *min* for tie-breaking, or the *leximax* criterion, allow for a better discrimination than the pure pessimistic criterion.

The interest of this framework is that we can define different degrees of contradiction, depending of the chosen underlying logic. We will only give an example of such instanciation here for illustrating the definition (see [40] for other examples).

We focus here on the LP_m logic as defined in [59]. This choice is mainly motivated by the fact that this logic is simple enough and has an inference relation that coincides with classical entailment whenever the knowledgebase is classically consistent (this feature is not shared by many paraconsistent logics).

- The language of LP_m is built up from the connectives $\wedge$, $\vee$, $\neg$, $\rightarrow$ and the constants $\top$, $\bot$.
- An interpretation ω for LP_m maps each atom to one of the three "truth values" *false*, *both*, *true*, the third truth value *both* meaning intuitively "both true and false". 3^{PS} is the set of all interpretations for LP_m. "Truth values" are ordered as follows: $false <_t both <_t true$.
  - $M(\top) = true$, $M(\bot) = false$
  - $M(\neg\alpha) = both$ iff $M(\alpha) = both$
    $M(\neg\alpha) = true$ iff $M(\alpha) = false$
  - $M(\alpha \wedge \beta) = \min_{\leq_t}(M(\alpha), M(\beta))$
  - $M(\alpha \vee \beta) = \max_{\leq_t}(M(\alpha), M(\beta))$
  - $M(\alpha \rightarrow \beta) = \begin{cases} true & \text{if } M(\alpha) = false \\ M(\beta) & \text{otherwise} \end{cases}$
- The set of models of a formula α is $\mathsf{Models}_{LP}(\alpha) = \{M \in 3^{PS} \mid M(\alpha) \in \{true, both\}\}$.
  Define $M! = \{x \in PS \mid M(x) = both\}$.
  Then $\mathsf{min}(\mathsf{Models}_{LP}(\alpha)) = \{M \in \mathsf{Models}_{LP}(\alpha) \mid \nexists M' \in \mathsf{Models}_{LP}(\alpha)$ s.t. $M'! \subset M!\}$.
  The consequence relation is defined by $\Delta \models_{LP_m} \alpha$ iff $min(\mathsf{Models}_{LP}(\Delta)) \subseteq \mathsf{Models}_{LP}(\alpha)$.
- The definitions of $A_{LP_m}(\Delta, \alpha)$ and $C_{LP_m}(\Delta, \alpha)$ are those by default; $C_{LP_m}(\Delta, \alpha)$ holds only if Δ has no classical model.

We have also to define the revision operator. Actually, the issue of revision in paraconsistent logic has never been considered so far. Expansion is not satisfactory as a revision operator for LP_m because it does not enable the purification task when Δ has no classical model ω (i.e., such that $\omega(x) \neq both$ for each $x \in PS$), whatever the test context. Among the many possible choices, we have considered the following revision operator, defined model-theoretically (for the sake of brevity, we characterize only its restriction to the case the revision formula α is a literal l).

Let $force(M, l)$ be the interpretation of 3^{PS} defined by (for every literal $l = x$ or $\neg x$):

$$\begin{cases} force(M, x)(x) = true \\ \forall y \in PS, y \neq x, force(M, x)(y) = M(y) \end{cases}$$

$$\begin{cases} force(M, \neg x)(x) = false \\ \forall y \in PS, y \neq x, force(M, x)(y) = M(y) \end{cases}$$

Then the revision operator is defined by:

$$\mathsf{Models}_{LP}(\Delta \star l) = \begin{cases} \{M \models \Delta \mid M(l) = true\} \text{ if this set is non-empty,} \\ \{force(M, l) \mid M \models \Delta \text{ and } M(l) = both\}, \text{ otherwise.} \end{cases}$$

Example 20. Given the standard atomic test context, we have:

- $d_C(\{\top\}) = 0$
- $d_C(\{a\}) = 0$
- $d_C(\{a \wedge b\}) = 0$
- $d_C(\{a \wedge b \wedge \neg a\}) = 1$
- $d_C(\{a \wedge b \wedge \neg a \wedge \neg b\}) = 2$

Let us see the result on a more complex example:

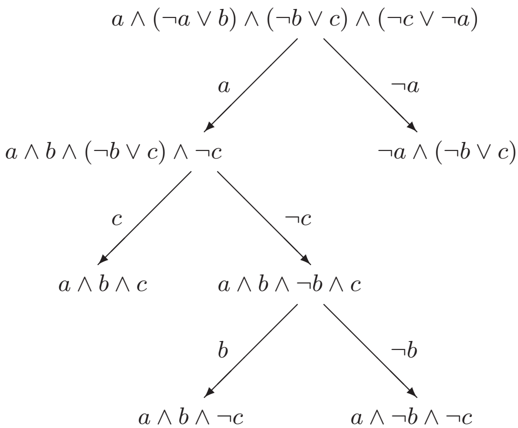

Fig. 1. Degree of contradiction in LP_m

Example 21. Let us consider the base $\varDelta = \{a \wedge (\neg a \vee b) \wedge (\neg b \vee c) \wedge (\neg c \vee \neg a)\}$. Figure 1 reports a plan of minimal cost (given the standard atomic context) which purifies $\varDelta$. So the degree of contradiction of $\varDelta$ is 3.

Note that what we define here is a measure of contradiction. To obtain a measure of coherence, as in the other sections, it is enough to define $I^{\mathcal{L}}_{klm_{(LP_m)}}(\varDelta) = |\mathcal{L}| - d_C(\varDelta)$ while considering the standard atomic test context.

Proposition 6. *Every knowledgebase that has a model $\in 3^{PS}$ has a finite degree of contradiction given any atomic or universal test context.*

Another example of instantiation with a paraconsistent logic can be obtained by taking the QC logic (see Section 8.2) as underlying logic. The revision operator in this case is the same one than for LP_m.

So this approach is highly configurable (underlying logic, test context), leading to several different particular measures of contradiction. It is less syntax-sensitive than the approaches in the previous sections, since a set of formulae is considered exactly as the conjunction of those formulae. This can be an advantage or a drawback, depending on the intended application. But the main advantage is that it does not trivialize when facing a single contradictory formula.

8 Model-Theoretic Analyses

Arguably, the most important logical language to analyse is that of classical logic. So far in this review we have considered a variety of approaches. Syntactic analysis is an obvious starting point for use with classical formulae. Useful alternatives for analysing classical formulae, which we have considered in previous sections, are based on information theory analysis, probability theory, and epistemic actions. Another alternative, which we consider in this section, is model-theoretic analysis. In this, inconsistent information is analysed in terms of the models of the information. Obviously this is not possible using classical models, because there is no model of a set of inconsistent formulae. To address this, we can use a three or four valued semantics, where one of the truth values denotes "inconsistency". In this section, we briefly consider an approach based on three-valued logic, and then review a framework based on quasi-classical logic which uses a four-valued semantics. Both of them can be used with a set of classical formulae.

8.1 Analysis of Three-Valued Models

In the proposal for three-valued models by [47], and a similar proposal by [26], a 3-interpretation is a truth assignment into {true,false} that does not map both a literal and its complement into false. This is extended to clauses so that a 3-interpretation satisfies a clause if and only if it satisfies some of the literals in the clause.

Example 22. For the set of formulae $\{\alpha, \neg\alpha \vee \neg\beta, \beta\}$, there are three 3-interpretations that satisfy: (X_1) $\alpha, \neg\alpha, \beta$ are true and $\neg\beta$ is false; (X_2) $\alpha, \beta, \neg\beta$ are true and $\neg\alpha$ is false; and (X_3) $\alpha, \neg\alpha, \beta, \neg\beta$ are true.

As shown by Grant [26], the 3-interpretations for a set of formulae can be analysed to obtained a degree of inconsistency. For this, the functions CCount and ICount are introduced. For a 3 interpretation X, $\mathsf{CCount}(X)$ gives the number of atoms in X for which either the atom or its complement is assigned false, and $\mathsf{ICount}(X)$ gives the number of atoms in X for which both the atom and its complement are assigned true. So CCount gives the number of atoms that are regarded as consistent, and ICount gives the number of atoms that are regarded as inconsistent, for X. In addition, LCount gives the total number of literals that are assigned by X. So $\mathsf{LCount}(X) = \mathsf{CCount}(X) + (2 \times \mathsf{ICount}(X))$.

Definition 17. *The degree of inconsistency for a 3-interpretation X (in the finite case) is the ratio*

$$Inc_G(X) = \frac{\mathsf{CCount}(X)}{\mathsf{LCount}(X)}$$

Example 23. So, on the previous example, we have $Inc_G(X_1) = 1/3$, $Inc_G(X_2) = 1/3$ and $Inc_G(X_3) = 0$.

This definition has been generalized by Grant to deal with countable 3-interpretations, and to take into account the domain for the 3-interpretations. A problem with the proposal is that it takes into account "too many models". Consider the small knowledgebase in Example 22 for which there are three 3-interpretations. With quasi-classical (QC) logic reviewed in the next subsection, only the third 3-interpretation would be considered a model. The intuition behind this is that if we consider disjunctive syllogism, or equivalently, the resolution proof rule, as being applicable here, then neither the first 3-interpretation nor the second 3-interpretation would be valid models for this set of formulae.

As we shall see below, QC logic has a more constrained semantics and proof theory resulting in a "more appropriate" selection of models, and as a consequence, a measure of consistency can be defined with some useful properties. However, when we consider analysing first-order QC models, we will see how Grant's degree of inconsistency can also be harnessed.

8.2 Analysis of Quasi-classical Models

Quasi-classical (QC) logic, a form of paraconsistent logic can be used as the basis of a framework, to measure inconsistency [31, 32, 33]. In this, each inconsistent set of formulae is reflected in the quasi-classical models for the set, and then the inconsistency is measured in the models.

Review of Propositional QC Logic. We review the propositional version of quasi-classical logic (QC Logic) [9, 29]. The language of propositional QC logic is that of classical propositional logic.

Let α be an atom, and let $\sim$ be a complementation operation such that $\sim\alpha$ is $\neg\alpha$ and $\sim(\neg\alpha)$ is α. The $\sim$ operator is not part of the object language, but it makes some definitions clearer.

Definition 18. *Let $\alpha_1 \vee .. \vee \alpha_n$ be a clause that includes a disjunct α_i and $n > 1$. The* **focus** *of $\alpha_1 \vee .. \vee \alpha_n$ by α_i, denoted $\otimes(\alpha_1 \vee .. \vee \alpha_n, \alpha_i)$, is defined as the clause obtained by removing α_i from $\alpha_1 \vee .. \vee \alpha_n$.*

Example 24. Let $\alpha \vee \beta \vee \gamma$ be a clause where α, β, and γ are literals. Hence, $\otimes(\alpha \vee \beta \vee \gamma, \beta) = \alpha \vee \gamma$.

Focus is used to capture a form of resolution in the semantics of QC logic. A model in QC logic is a form of Herbrand model.

Definition 19. *Let $\mathcal{A}$ be a set of atoms. Let $\mathcal{O} = \{+\alpha \mid \alpha \in \mathcal{A}\} \cup \{-\alpha \mid \alpha \in \mathcal{A}\}$ be the set of objects defined as follows, where $+\alpha$ is a positive object, and $-\alpha$ is a negative object. We call any $X \in \wp(\mathcal{O})$ a* **QC model***. So X can contain both $+\alpha$ and $-\alpha$ for some atom α.*

For each atom $\alpha \in \mathcal{L}$, and each $X \in \wp(\mathcal{O})$, $+\alpha \in X$ means that in X there is **a reason for** the belief α and $-\alpha \in X$ means that in X there is **a reason for** the belief $\neg\alpha$. This effectively gives us a four-valued semantics. Though for non-atomic formulae the semantics, defined next, is significantly different to [6].

Definition 20. *Let $\models_s$ be a satisfiability relation called* **strong satisfaction**. *For a model X, we define $\models_s$ as follows, where $\alpha_1, ..., \alpha_n$ are literals in $\mathcal{L}$, $n > 1$, and α is a literal in $\mathcal{L}$.*

$$X \models_s \alpha \textit{ iff there is a reason for the belief } \alpha \textit{ in } X$$

$$\begin{array}{l} X \models_s \alpha_1 \vee ... \vee \alpha_n \\ \quad \textit{iff } [X \models_s \alpha_1 \textit{ or ... or } X \models_s \alpha_n] \\ \qquad \textit{and } \forall i \textit{ s.t. } 1 \leq i \leq n \\ \qquad\quad [X \models_s \sim\alpha_i \textit{ implies } X \models_s \otimes(\alpha_1 \vee ... \vee \alpha_n, \alpha_i)] \end{array}$$

For $\alpha, \beta, \gamma \in \mathcal{L}$, we extend the definition as follows,

$$\begin{array}{l} X \models_s \alpha \wedge \beta \textit{ iff } X \models_s \alpha \textit{ and } X \models_s \beta \\ X \models_s \neg\neg\alpha \vee \gamma \textit{ iff } X \models_s \alpha \vee \gamma \\ X \models_s \neg(\alpha \wedge \beta) \vee \gamma \textit{ iff } X \models_s \neg\alpha \vee \neg\beta \vee \gamma \\ X \models_s \neg(\alpha \vee \beta) \vee \gamma \textit{ iff } X \models_s (\neg\alpha \wedge \neg\beta) \vee \gamma \\ X \models_s \alpha \vee (\beta \wedge \gamma) \textit{ iff } X \models_s (\alpha \vee \beta) \wedge (\alpha \vee \gamma) \\ X \models_s \alpha \wedge (\beta \vee \gamma) \textit{ iff } X \models_s (\alpha \wedge \beta) \vee (\alpha \wedge \gamma) \end{array}$$

Definition 21. *For $X \in \wp(\mathcal{O})$ and $\Delta \in \wp(\mathcal{L})$, let $X \models_s \Delta$ denote that $X \models_s \alpha$ holds for every α in Δ. Let $\mathsf{QC}(\Delta) = \{X \in \wp(\mathcal{O}) \mid X \models_s \Delta\}$ be the set of QC models for Δ.*

A key feature of the QC semantics is that there is a model for any formula, and for any set of formulae.

Example 25. Let $\Delta = \{\neg\alpha \vee \neg\beta \vee \gamma, \neg\alpha \vee \gamma, \neg\gamma\}$, where $\alpha, \beta, \gamma \in \mathcal{A}$, and let $X = \{-\alpha, -\beta, -\gamma\}$. So $X \models_s \neg\alpha$, $X \models_s \neg\beta$ and $X \models_s \neg\gamma$. Also, $X \models_s \sim\gamma$. Hence, $X \models_s \neg\alpha \vee \gamma$, and $X \models_s \neg\alpha \vee \neg\beta$, and so, $X \models_s \neg\alpha \vee \neg\beta \vee \gamma$. Hence every formula in Δ is strongly satisfiable in X.

Strong satisfaction is used to define a notion of entailment for QC logic. There is also a natural deduction proof theory for propositional QC logic [29] and a semantic tableau version for first-order QC logic [30]. Entailment for QC logic for propositional CNF formulae is coNP-complete, and via a linear time transformation these formulae can be handled using classical logic theorem provers [53].

The definitions for QC models and for strong satisfaction provide us with the basic concepts for measuring inconsistency. QC logic exhibits the nice feature that no attention needs to be paid to a special form that the formulae in a set of premises should have. This is in contrast with other paraconsistent logics where two formulae identical by definition of a connective in classical logic may not yield the same set of conclusions. For example, in QC logic, β is entailed by both $\{(\neg\alpha \rightarrow \beta), \neg\alpha\}$ and $\{\alpha \vee \beta, \neg\alpha\}$ and γ is entailed by $\{\gamma \wedge \neg\gamma\}$ and $\{\gamma, \neg\gamma\}$. QC logic is much better behaved in this respect than other paraconsistent logics such as C_ω [17], and consistency-based logics such as [7]. Furthermore, the semantics of QC logic directly models inconsistent sets of formulae.

Definition 22. *Let $\Delta \in \wp(\mathcal{L})$. Let $\mathsf{MQC}(\Delta) \subseteq \mathsf{QC}(\Delta)$ be the set of minimal QC models for Δ, defined as follows:*

$$\mathsf{MQC}(\Delta) = \{X \in \mathsf{QC}(\Delta) \mid \textit{ if } Y \subset X, \textit{ then } Y \notin \mathsf{QC}(\Delta)\}$$

Example 26. Consider the following sets of formulae.

$$\begin{aligned}
&\mathsf{MQC}(\{\alpha \wedge \neg\alpha, \alpha \vee \beta, \neg\alpha \vee \gamma\}) \\
&\qquad\qquad = \{\{+\alpha, -\alpha, +\beta, +\gamma\}\} \\
&\mathsf{MQC}(\{\neg\alpha \wedge \alpha, \beta \vee \gamma\}) \\
&\qquad\qquad = \{\{+\alpha, -\alpha, +\beta\}, \{+\alpha, -\alpha, +\gamma\}\} \\
&\mathsf{MQC}(\{\alpha \vee \beta, \neg\alpha \vee \gamma\}) \\
&\qquad\qquad = \{\{+\beta, +\gamma\}, \{+\alpha, +\gamma\}, \{-\alpha, +\beta\}\}
\end{aligned}$$

Whilst four-valued logic [6] also directly models inconsistent sets of formulae, there are too many Belnap models in many situations. Consider for example $\{\alpha \vee \beta, \neg\alpha\}$. There is one minimal QC model $\{-\alpha, +\beta\}$, but there are a number of Belnap models that satisfy this set. QC logic has a reduced number of models because of the constraint in the definition of strong satisfaction for disjunction that ensures that if the complement of a disjunct holds in the model, then the resolvent should also hold in the model. This strong constraint means that various other proposals for many-valued logic will tend to have more models for any given knowledgebase than QC logic. This increases the number of models that need to be analysed and it underspecifies the nature of the conflicts. These shortcomings of Belnap's four-valued logic also apply to three-valued logics such as 3-interpretations by [47], and a similar proposal by [26].

Measuring Coherence of QC Models. We now consider a measure of inconsistency called coherence [31]. The opinionbase of a QC model X is the set of atomic beliefs (atoms) for which there are reasons for or against in X, and the conflictbase of X is the set of atomic beliefs with reasons for and against in X.

Definition 23. *Let $X \in \wp(\mathcal{O})$.*

$$\begin{aligned}
\mathsf{Conflictbase}(X) &= \{\alpha \mid +\alpha \in X \textit{ and } -\alpha \in X\} \\
\mathsf{Opinionbase}(X) &= \{\alpha \mid +\alpha \in X \textit{ or } -\alpha \in X\}
\end{aligned}$$

In finding the minimal QC models for a set of formulae, minimization of each model forces minimization of the conflictbase of each model. As a result of this minimization, if $\Delta \in \wp(\mathcal{L})$, and $X, Y \in \mathsf{MQC}(\Delta)$, then $\mathsf{Conflictbase}(X) = \mathsf{Conflictbase}(Y)$.

Increasing the size of the conflictbase, with respect to the size of the opinionbase, decreases the degree of coherence, as defined below.

Definition 24. *The* Coherence *function from* $\wp(\mathcal{O})$ *into* $[0,1]$*, is given below when* X *is non-empty, and* $\mathsf{Coherence}(\emptyset) = 1$*.*

$$\mathsf{Coherence}(X) = 1 - \frac{|\mathsf{Conflictbase}(X)|}{|\mathsf{Opinionbase}(X)|}$$

If $\mathsf{Coherence}(X) = 1$, then X is a totally coherent, and if $\mathsf{Coherence}(X) = 0$, then X is totally incoherent, otherwise, X is partially coherent/incoherent.

Example 27. Let $X \in \mathsf{MQC}(\{\neg\alpha \wedge \alpha, \beta \wedge \neg\beta, \gamma \wedge \neg\gamma\})$, $Y \in \mathsf{MQC}(\{\alpha, \neg\alpha \vee \neg\beta, \beta, \gamma\})$, and $Z \in \mathsf{MQC}(\{\neg\alpha, \beta, \neg\neg\gamma \wedge \gamma\})$. So $\mathsf{Coherence}(X) = 0$, $\mathsf{Coherence}(Y) = 1/3$, and $\mathsf{Coherence}(Z) = 2/3$.

Different minimal QC models for the same knowledgebase are not necessarily equally coherent, since different models for the same knowledgebase may have different opinionbases, though they will have the same conflictbase.

Example 28. Let $\Delta = \{\alpha, \neg\alpha, \beta \vee \gamma, \beta \vee \delta\}$, and let $X = \{+\alpha, -\alpha, +\beta\}$ and $Y = \{+\alpha, -\alpha, +\gamma, +\delta\}$. So $\mathsf{MQC}(\Delta) = \{X, Y\}$, and $\mathsf{Coherence}(X) = 1/2$ and $\mathsf{Coherence}(Y) = 2/3$.

We extend coherence to knowledgebases as follows.

Definition 25. *Let* $\Delta \in \wp(\mathcal{L})$*. Assign* $\mathsf{Coherence}(\Delta)$ *the maximum value in* $\{\mathsf{Coherence}(X) \mid X \in \mathsf{MQC}(\Delta)\}$*.*

Example 29. Let $\Delta = \{\phi \wedge \neg\phi, \alpha \vee (\beta \wedge \gamma \wedge \delta)\}$ and $\Delta' = \{\phi \wedge \neg\phi, (\alpha \wedge \beta) \vee (\gamma \wedge \delta)\}$. Also let $X_1 = \{+\phi, -\phi, +\alpha\}$, $X_2 = \{+\phi, -\phi, +\beta, +\gamma, +\delta\}$, $Y_1 = \{+\phi, -\phi, +\alpha, +\beta\}$, and $Y_2 = \{+\phi, -\phi, +\gamma, +\delta\}$. So, $\mathsf{MQC}(\Delta) = \{X_1, X_2\}$ and $\mathsf{MQC}(\Delta') = \{Y_1, Y_2\}$. Also, $\mathsf{Coherence}(X_1) = 1/2$, $\mathsf{Coherence}(X_2) = 3/4$, $\mathsf{Coherence}(Y_1) = 2/3$, and $\mathsf{Coherence}(Y_2) = 2/3$. So $\mathsf{Coherence}(\Delta) > \mathsf{Coherence}(\Delta')$.

Note that the definition of the coherence of a knowledgebase is an optimistic one, since it is based on the maximal coherence value of its models. But taking other aggregation functions could be interesting. For example taking a leximax function would allow for more discrimination. And taking the minimum or a mean can lead to other interesting measures. Such generalisations have not been considered yet, but can be a starting point for further work.

Significance Functions. The QC logic framework for measuring inconsistency has been extended to measuring the significance of inconsistencies arising in QC models, and thereby in sets of formulae [32]. The approach is based on specifying the relative significance of incoherent models using additional information, encoded as a mass assignment, which is defined below.

Definition 26. *A* **mass assignment** m *for* $\mathcal{O}$ *is a function from* $\wp(\mathcal{O})$ *into* $[0,1]$ *such that:*

(1) *If* $X \subseteq \mathcal{O}$ *and* $\mathsf{Coherence}(X) = 1$, *then* $m(X) = 0$

(2) $\Sigma_{X \subseteq \mathcal{O}}\, m(X) = 1$

Condition 1 ensures mass is only assigned to models that contain conflicts and condition 2 ensures the total mass distributed sums to 1. A mass assignment can be localized on small subsets of $\mathcal{O}$, spread over many subsets of $\mathcal{O}$, or limited to large subsets of $\mathcal{O}$. A mass assignment can be regarded as a form of metaknowledge, and so it needs to be specified for an application area, where the application area is characterized by $\mathcal{O}$.

Example 30. Let $\mathcal{O} = \{+\alpha, -\alpha, +\beta, -\beta\}$. A mass assignment m is given by $m(\{+\alpha, -\alpha\}) = 0.2$ and $m(\{+\beta, -\beta\} = 0.8$. Another mass assignment m' is $m'(\{+\alpha, -\alpha\}) = 0.2$, $m'(\{+\alpha, -\alpha, -\beta\}) = 0.6$, and $m'(\{+\alpha, -\alpha, +\beta, -\beta\}) = 0.2$.

A significance function gives an evaluation of the significance of the conflicts in a QC model. This evaluation is in the range $[0, 1]$ with 0 as least significant and 1 as most significant.

Definition 27. *Let m be a mass assignment for $\mathcal{O}$. A* **significance function** *for $\mathcal{O}$, denoted S, is a function ¿from $\wp(\mathcal{O})$ into $[0, 1]$. A* **mass-based significance function** *for m, denoted S^m, is a significance function defined as follows for each $X \in \wp(\mathcal{O})$.*

$$S^m(X) = \Sigma_{Y \subseteq X} m(Y)$$

The definitions for mass assignment and mass-based significance correspond to mass assignment and belief functions (respectively) in Dempster-Shafer theory [66]. However, here they are used to formalise significance rather than uncertainty.

Proposition 7. *Let m be a mass assignment for $\mathcal{O}$. If S^m is a significance function, then the following property of simple cumulativity holds for all $X, Y \in \wp(\mathcal{O})$: $X \subseteq Y$ implies $S^m(X) \leq S^m(Y)$.*

Given that simple cumulativity holds, we see that specifying significance in terms of mass assignment is more efficient than directly specifying the significance.

Proposition 8. *Let m be a mass assignment for $\mathcal{O}$. Let S^m be a mass-based significance function. For all $X, Y \in \wp(\mathcal{O})$,*

$$\begin{array}{l} (1)\ S^m(X \cup Y) \geq (S^m(X) + S^m(Y) - S^m(X \cap Y)) \\ (2)\ S^m(X) + S^m(X^c) \leq 1 \end{array}$$

So mass-based significance is not additive. Also the remaining significance need not be for the complement of X (ie, X^c). Some may be assigned to models not disjoint from X. We now consider some constraints on mass assignments that give useful properties for mass-based significance.

Definition 28. *Let m be a mass assignment for $\mathcal{O}$. m is* **focal** *iff for all $X \in \wp(\mathcal{O})$ $m(X) \geq 0$ when $\mathsf{Coherence}(X) = 0$ and $m(X) = 0$ when $\mathsf{Coherence}(X) > 0$. m is* **solo** *iff for all $\{+\alpha, -\alpha\} \in \wp(\mathcal{O})$ $m(\{+\alpha, -\alpha\}) \geq 0$ and for all other $X \in \wp(\mathcal{O})$, $m(X) = 0$.*

A focal mass assignment puts the mass onto the totally incoherent models, and a solo mass assignment puts the mass on the smallest totally incoherent models. For all m, if m is a solo mass assignment for $\mathcal{O}$, then m is focal mass assignment for $\mathcal{O}$. Significance is additive for totally incoherent models when the mass assignment is solo.

Proposition 9. *Let m be a solo mass assignment for $\mathcal{O}$. Let S^m be a mass-based significance function and let $X \in \wp(\mathcal{O})$. If $\mathsf{Coherence}(X) = 0$, then $S^m(X) + S^m(X^c) = 1$.*

A useful feature of a focal mass-based significance function is that as the number of conflicts rises in a model, then the significance of the model rises. This is formalized by the following notion of conflict cumulativity.

Proposition 10. *Let m be a focal mass assignment for $\mathcal{O}$. If S^m is a significance function, then the following property of conflict cumulativity holds for all $X, Y \in \wp(\mathcal{O})$: $\mathsf{Conflictbase}(X) \subseteq \mathsf{Conflictbase}(Y)$ implies $S^m(X) \leq S^m(Y)$.*

We now extend the significance functions to knowledgebases. Since $\mathsf{MQC}(\Delta)$ is not necessarily a singleton, the significance for a set of formulae Δ is the lowest significance obtained for an $X \in \mathsf{MQC}(\Delta)$. This means we treat the information in Δ as a "disjunction" of QC models, and we regard each of those models as equally acceptable, or equivalently we regard each of those models as equally representative of the information in Δ. As with Definition 25, the following definition is an optimistic view, in the sense that taking the higher coherence value and lower significance value is better.

Definition 29. *Let $\Delta \in \wp(\mathcal{L})$. We extend the definition for a significance function S^m to knowledgebases as follows:*

$$S^m(\Delta) = \mathsf{min}(\{S^m(X) \mid X \in \mathsf{MQC}(\Delta)\})$$

Some knowledgebases have zero significance. Clearly, if $\Delta \not\vdash \bot$, then $S^m(\Delta) = 0$.

Example 31. Let $\Omega = \{+\alpha, -\alpha, +\beta, -\beta, +\gamma, -\gamma\}$. Let $m(\{+\alpha, -\alpha, +\beta\}) = 0.3$, $m(\{+\alpha, -\alpha\}) = 0.6$, and $m(\{+\beta, -\beta, +\gamma\}) = 0.1$. So $S^m(\{\alpha \wedge \neg\alpha, \beta \vee \gamma\}) = 0.6$

In order to determine the set $\mathcal{O}$ for which a mass function is defined, we can use the delineation function as follows.

Definition 30. *For $\Delta \in \wp(\mathcal{L})$, $\mathsf{Delineation}(\Delta) = \{+\alpha, -\alpha \mid \alpha \in \mathsf{Atoms}(\Delta)\}$.*

Example 32. Let $\Delta_1 = \{\neg\alpha, \alpha \vee \beta, \neg\beta\}$, $\Delta_2 = \{\alpha \vee \beta, \neg\alpha \wedge \alpha\}$, and $\Delta_3 = \{\beta, \neg\alpha \vee \neg\beta\}$. Let $\mathcal{O} = \mathsf{Delineation}(\Delta_1 \cup \Delta_2 \cup \Delta_3) = \{+\alpha, -\alpha, +\beta, -\beta\}$. Also let $m(\{+\alpha, -\alpha, +\beta, -\beta\}) = 0.2$ and $m(\{+\alpha, -\alpha\}) = 0.8$. So $S^m(\Delta_1) = 1$, $S^m(\Delta_2) = 0.8$, and $S^m(\Delta_3) = 0$.

The next result captures a notion of monotonicity for mass-based significance.

Proposition 11. *Let $\Delta \in \wp(\mathcal{L})$ and $\alpha \in \mathcal{L}$. Let m be a mass assignment for* $\mathsf{Delineation}(\Delta \cup \{\alpha\})$. *If S^m is a significance function, then $S^m(\Delta) \leq S^m(\Delta \cup \{\alpha\})$.*

Another approach to analysing the significance of inconsistency is possibility theory [19]. Let (ϕ, α) be a weighted formula where ϕ is a classical formula and $\alpha \in [0, 1]$. A possibilistic knowledgebase B is a set of weighted formulae. An α-cut of a possibilistic knowledgebase, denoted $B_{\geq\alpha}$, is $\{(\psi, \beta) \in B \mid \beta \geq \alpha\}$. The inconsistency degree of B, denoted $Inc(B)$, is the maximum value of α such that the α-cut is inconsistent. Possibility theory can also be used to extend classical logic, so that the proof rules propagate the possibility weights. This logic is called possibilistic logic and it offers complementary reasoning to that offered by QC logic.

Possibilistic logic and QC logic can be combined to give quasi-possibilistic logic [18]. This combined logic can handle plain conflicts taking place at the same level of certainty, as in QC logic, and take advantage of the stratification of the knowledgebase into certainty layers for introducing gradedness in conflict analysis, as in possibilistic logic. Moreover, quasi-possibilistic logic can be used to generalize the QC logic framework for measuring the degree and significance of inconsistencies.

Compromising on Inconsistency. In the following, we define the compromise relation to prefer knowledgebases with models with a greater opinionbase and a smaller conflictbase.

Definition 31. *Let $\Delta, \Delta' \in \wp(\mathcal{L})$. The* **compromise relation**, *denoted $\preceq$, is defined as follows:*

$$\begin{aligned} \Delta \preceq \Delta' \ \textit{iff}\ & \forall X \in \mathsf{MQC}(\Delta) \ \textit{and}\ \exists Y \in \mathsf{MQC}(\Delta') \\ & \textit{such that}\ \mathsf{Conflictbase}(X) \subseteq \mathsf{Conflictbase}(Y) \\ & \textit{and}\ \mathsf{Opinionbase}(Y) \subseteq \mathsf{Opinionbase}(X) \end{aligned}$$

We read $\Delta \preceq \Delta'$ as Δ is a preferred compromise to Δ'. Let $\Delta \prec \Delta'$ denote $\Delta \preceq \Delta'$ and $\Delta' \not\preceq \Delta$. Also let $\Delta \simeq \Delta'$ denote $\Delta \preceq \Delta'$ and $\Delta' \preceq \Delta$.

Example 33. If $\Delta = \{\alpha \wedge \beta \wedge \gamma\}$, and $\Delta' = \{\alpha \wedge \neg\alpha, \beta \vee \gamma\}$, then $\Delta \prec \Delta'$, since the following hold,

$$\begin{gathered} \mathsf{MQC}(\Delta) = \{\{+\alpha, +\beta, +\gamma\}\} \\ \mathsf{MQC}(\Delta') = \{\{+\alpha, -\alpha, +\beta\}, \{+\alpha, -\alpha, +\gamma\}\} \end{gathered}$$

Example 34. If $\Delta = \{\alpha \wedge \neg\alpha \wedge \beta\}$ and $\Delta' = \{\beta\}$, then $\Delta \not\preceq \Delta'$, and $\Delta' \not\preceq \Delta$, since $\mathsf{MQC}(\Delta) = \{\{+\alpha, -\alpha, +\beta\}\}$ and $\mathsf{MQC}(\Delta') = \{\{+\beta\}\}$. Though $\mathsf{Coherence}(\Delta) < \mathsf{Coherence}(\Delta')$.

Example 35. If $\Delta = \{\alpha \vee \beta\}$ and $\Delta' = \{\alpha \vee \gamma\}$, then $\Delta \not\preceq \Delta'$, and $\Delta' \not\preceq \Delta$, since $\mathsf{MQC}(\Delta) = \{\{+\alpha\}, \{+\beta\}\}$ and $\mathsf{MQC}(\Delta') = \{\{+\alpha\}, \{+\gamma\}\}$. Though $\mathsf{Coherence}(\Delta) = \mathsf{Coherence}(\Delta')$.

We now motivate the compromise relation. For checking whether $\Delta \preceq \Delta'$ holds, we want to compare the minimal QC models of Δ with the minimal QC models of Δ'. First, we want each minimal QC model of Δ to have a conflictbase that is a subset of the conflictbase of each minimal QC model of Δ'. Second, we want for each minimal QC model X of Δ, for there to be a minimal QC model Y of Δ' such that the opinionbase of Y is a subset of the opinionbase of X. This is to ensure that Δ is not less conflicting than Δ' because Δ has less information in it. The reason we use the condition $\mathsf{Opinionbase}(Y) \subseteq \mathsf{Opinionbase}(X)$ rather than $Y \subseteq X$ is that if Y is more conflicting than X, then this will be reflected in the membership of Y but not in the membership of $\mathsf{Opinionbase}(Y)$. The reason we only seek one minimal QC model of Δ' for the comparison with all the minimal QC models of Δ is so that we can handle disjunction in Δ' as illustrated by Example 33. Useful properties of the compromise relation include it is a pre-order relation and it is syntax independent.

Let us note that, although the compromise relation and coherence function are logically independent notions, they are "philosophically" related, since in both case it is better when the conflicts decrease or when the information increase.

Measuring First-Order Inconsistency. Using the first-order version of quasi-classical logic [30], the QC logic framework for measuring inconsistency has been extended to first-order logic [33]. In first-order QC logic, the strong satisfaction relation is extended for universal and existential quantification.

- A QC model M, with a variable assignment A, satisfies a formula $\exists X \alpha$ if and only M satisfies α with some variable assignment A' that differs from A in at most the assignment for X.
- A QC model M, with a variable assignment A, satisfies a formula $\forall X \alpha$ if and only if M satisfies α with all variable assignments A' that differ from A in at most the assignment for X.

As with propositional QC logic, the models are a form of Herbrand model. The definitions for minimal QC model, for coherence, and for compromise relation can be used with first-order information.

In another development, the degree of inconsistency as presented by Grant [26], has been incorporated into the QC logic framework for measuring inconsistency in first-order information [25]. In this, both the language and the domain is taken into account. In the following, we restrict the presentation to a first-order language with constant symbols and no function symbols.

Definition 32. *For a language $\mathcal{L} = \langle \mathcal{P}, \mathcal{C} \rangle$, where $\mathcal{P}$ is a set of predicates represented in the form $P(n)$, with P being the predicate symbol and n being the arity of the predicate, $\mathcal{C}$ is a set of constants, and D is a domain,*

$$\mathsf{Groundatoms}(\mathcal{L}, D) = \{P(c_1, .., c_n) \mid P(n) \in \mathcal{P} \textit{ and } c_1, .., c_n \in D\}$$

This is used for a measure as a ratio between 0 and 1 whose denominator is the total possible number of inconsistencies in the bistructure.

Definition 33. *The measure of inconsistency for a model M in the context of a language $\mathcal{L}$ and a domain D is given by the* $\mathsf{ModelInc}$ *function giving a value in* $[0, 1]$ *as follows.*

$$\mathsf{ModelInc}(M, \mathcal{L}, D) = \frac{|\mathsf{Conflictbase}(M)|}{|\mathsf{Groundatoms}(\mathcal{L}, D)|}$$

Example 36. Let $\mathcal{L} = \langle\{P(2), R(1)\}\{\}\rangle$. Hence, P is a binary predicate and R is a monadic predicate. Let $D = \{a, b, c\}$, and $M = \{+P(a, a), -P(a, a), +R(a), -R(b), +P(b, c)\}$, $|\mathsf{Groundatoms}(\mathcal{L}, D)| = 12$ (9 ground atoms for P and 3 for R). $\mathsf{Conflictbase}(M) = \{P(a, a)\}$. Hence, $\mathsf{ModelInc}(M, \mathcal{L}, D) = \frac{1}{12}$.

The $\mathsf{ModelInc}$ definition provides the basis of a richer framework for comparing first-order formulae. In the following example, we compare some inconsistent formulae. For this, we consider the preferred QC models: These are the minimal QC models with a minimal conflictbase. Given a language $\mathcal{L}$ and a domain D, the value of $\mathsf{ModelInc}$ is the same for all preferred QC models for a knowledgebase.

Example 37. Let $\mathcal{L} = \langle\{P(2)\}, \{\}\rangle$ and $D = \{a, b, c\}$.

1. $\Delta_1 = \{\forall x \forall y (P(x, y) \wedge \neg P(x, y))\}$ has one preferred QC model which is represented by $M_1 = \{+P(a, a), -P(a, a), \ldots, +P(c, c), -P(c, c)\}$, so $\mathsf{ModelInc}(M_1, \mathcal{L}, D) = \frac{9}{9} = 1$. M_1 is totally inconsistent.
2. $\Delta_2 = \{\exists x \exists y (P(x, y) \wedge \neg P(x, y))\}$ has 9 preferred QC models. One of them is $M_{21} = \{-P(a, b),$ $+ P(a, b)\}$, so $\mathsf{ModelInc}(M_{21}, \mathcal{L}, D) = \frac{1}{9}$.
3. $\Delta_3 = \{\forall x \exists y (P(x, y) \wedge \neg P(x, y))\}$ has 9 preferred QC models. One is $M_{31} = \{+P(a, a),$ $- P(a, a), +P(b, c), -P(b, c), +P(c, a), -P(c, a)\}$, so $\mathsf{ModelInc}(M_{31}, \mathcal{L}, D) = \frac{3}{9} = \frac{1}{3}$.
4. $\Delta_4 = \{\exists x \forall y (P(x, y) \wedge \neg P(x, y))\}$ has 3 preferred QC models. One is $M_{41} = \{+P(b, a),$ $- P(b, a), +P(b, b), -P(b, b), +P(b, c), -P(b, c)\}$, so $\mathsf{ModelInc}(M_{41}, \mathcal{L}, D) = \frac{3}{9} = \frac{1}{3}$.

Comparing quantified formulae is potentially important in diverse applications such as analysing systems specifications and analysing sources of information as a precursor to selecting sources for merging. These applications potentially include consideration of information that violates integrity constraints. This framework incorporates a notion of quasi-equality, which is weaker than classical equality, but can be formalized as an extension to QC logic for reasoning about integrity constraint violations.

9 Choosing a Good Measure

In the previous sections we have presented the existing measures of contradiction and of information for (possibly) inconsistent information. We have tried to highlight the advantages and the typical uses of each of its measures. In this section, we will try to compare them in order to highlight their differences and to guide the choice of a particular measure.

9.1 Logical Properties

A very convenient way to compare several approaches to the same problem is to propose a set of logical properties, aiming at capturing the typical wanted behaviours, and to compare the approaches with respect to the properties satisfied/dissatisfied.

Setting these properties have several advantages: first, it allows to "abstract" the discussion, i.e. to drop the discussion from the examples that are the particular approaches, for a discussion on the wanted behaviour for the given problem. Second, it gives a mean to compare the different approaches and to highlight the differences of behaviour and underlying rationale (in a much more explicit way that when building examples that are correctly handled by one approach and badly by an other). Thirdly, this allows us to define in one shot a whole family of methods (the ones that satisfy a set of properties), instead of only one particular one. And in the case where there is only one approach satisfying a set of properties, then it usually gives a nice comprehensive definition of the approach.

Setting a set of properties have usually accelerated the development of a corresponding field. We can cite for example the work of Arrow for social choice theory (voting theory) [3], Savage for decision theory [64]. In artificial intelligence, the same happened for non-monotonic inference relations [51, 43, 46], belief revision [1, 24], and belief merging [62, 41, 42].

So, it might be of great interest to find a set of logical properties for information measures (that allows non-trivial information content for inconsistent information), and for contradiction measures.

Information Measures. Lozinskii [49] gives a set of properties that a measure of quantity of information should satisfy.

He stated those properties in first order logic. Recasted in propositional logic, those conditions can be summarized as follows. I is a function from $\mathcal{L}_{PS}$ to a numeric scale with least element 0 such that:

1. If $\Delta = \emptyset$, then $I(\Delta) = 0$
2. If $\Delta = \{a \in \mathcal{L}_{PS}\} \cup \{\neg a \mid a \in \mathcal{L}_{PS}\}$, then $I(\Delta) = 0$
3. If Δ is consistent, and α is a consequence of Δ, then $I(\Delta \cup \{\alpha\}) = I(\Delta)$
4. If $\Delta \cup \{\alpha\}$ is consistent and α is not a consequence of Δ, then $I(\Delta \cup \{\alpha\}) > I(\Delta)$
5. If Δ is consistent and α is a consequence of Δ, then $I(\Delta \cup \{\neg\alpha\}) < I(\Delta)$
6. If $\forall a \in \mathcal{L}_{PS}\ \Delta \vdash a$ or $\Delta \vdash \neg a$, then $\forall \Delta'\ I(\Delta) \geq I(\Delta')$

The first condition states that an empty knowledgebase contains no information. The second condition states that if a knowledgebase contains all the atoms and their negation of the language, it gives also no information. In the first condition it was caused by a lack of information, in the second one it is because of an overload of (contradictory) information. The third condition states that adding a consequence of a knowledgebase does not change anything of the information content. That implies, in particular, an irrelevance of syntax, since there is no difference between an explicit formula of Δ and an implicit one[4]. It says also that several occurrences of the same formula (or several way to derive the same formula), do not improve the information content. The fourth condition says that adding a (consistent) formula that is not a consequence of a consistent knowledgebase increases the amount of information. The fifth condition says that adding a (consistent) formula that contradicts a consistent knowledgebase decreases the amount of information. This condition relates the contradiction of a base and its information content. Thus, as expected, the introduction of a contradiction decreases the amount of information. The sixth condition states that a complete knowledgebase, thus having exactly one classical model, has the highest possible information content. This condition is quite natural, and quite close to the idea of Shannon's information theory.

It is not surprising that the information measure $I_l(\Delta)$ of section 5 satisfies those conditions.

Knight in [39, 38] propose also a set of logical properties for measures of quantity of information. Most of them are equivalent to Lozinskii's ones. The different ones are:

7. $0 \leq I(\Delta) \leq |PS|$
8. $I(\{\alpha\}) = 0$ if and only if $\vdash \alpha$ or $\vdash \neg\alpha$
9. If Δ and Δ' are logically equivalent[5], then $I(\Delta) = I(\Delta')$

Condition 7 simply puts bounds on the value of the information measure. Putting 0 as minimum is quite natural (and is already asked in Lozinskii's conditions). So the addition of this property is to put a maximum on the value. It is clear that in a finite setting, the information value must have an upper bound, but we are not sure that giving a precise bound is useful (taking any strictly increasing function of $|PS|$ would basically give the same thing, as acknowledge by Knight [39]). Condition 8 states that the only singleton knowledgebase having null information value are tautologies and contradictions. Condition 9 is an irrelevance of syntax condition, basically saying that we can exchange any formula of a knowledgebase by a logically equivalent one without changing the amount of information in the knowledgebase.

Those two last conditions are not similar to the previous ones. They basically say that information measures can cope with inconsistency only because we work

[4] An explicit formula of a knowledge base is a formula $\alpha \in \Delta$. An implicit formula of a knowledge base is a formula $\alpha \notin \Delta$ and $\Delta \vdash \alpha$.

[5] We say that two knowledgebases are logically equivalent if for each formula of a base, there is a formula in the other base that is logically equivalent to the first one.

with sets of formulae. But if one is faced with a unique inconsistent formula, it continues to have a null information value. So it is a very strong assumption that forbids consideration of inconsistent formulae. Whereas the information measures proposed by Knight and Lozinskii satisfy those two properties, the one proposed by Konieczny et al. [40] do not satisfy them, since they differentiate the information content of inconsistent formulae.

In addition, there are also two conditions on the relation between the language and the information value. The first one is given by Lozinskii in [48], the second one by Knight [39, 38].

9. If $PS \subseteq PS'$, then $max_{\Delta \in \mathcal{L}_{PS}}(I(\Delta)) \leq max_{\Delta' \in \mathcal{L}_{PS'}}(I(\Delta'))$
10. If $PS \subseteq PS'$, then $I_{\Delta \in \mathcal{L}_{PS}}(\Delta) = I_{\Delta \in \mathcal{L}_{PS'}}(\Delta)$

Condition 10 says that extending the language does not change the information content of a given knowledgebase. But, as expected, condition 9 says that extending the language allows to express more things, and so the upper bound of the information measure in the extended language is higher than the one in the original language.

One can think of other meaningful conditions, but the ones given here seem to be a good place to begin.

Contradiction Measures. The story is not the same for contradiction measures. It is much more difficult to state properties for contradiction measures than for information ones. Since classical logic is not the right tool for talking about inconsistency, it is difficult to state interesting logical properties using only classical logic.

In order to state the wanted properties, one should need to use a paraconsistent logic. But there are a lot of different paraconsistent logics (see e.g. [28, 13, 60]), so choosing a particular logic is already a real, non-trivial commitment. So one has to be careful for avoiding stating *ad hoc* properties, according to a given paraconsistent logic.

This maybe explains why there is not yet any proposal of such set of properties for contradiction measures. But this is an interesting and important open question.

9.2 Comparison of the Measures

Let us first talk about the information measures presented in this paper. For comparing Lozinskii's measure (section 5) and Knight's ones (section 6.2), Let us quote Knight [39]:

> "For when we look at proper subsets of Γ we fail to account for the affect the remaining sentences of Γ have on the sentences of the subset. Thus the author asserts the superiority of $I^{\mathcal{L}}_{k1}$ - and, indeed, $I^{\mathcal{L}}_{k2}$ - over such information measures as Lozinskii's [48, 49] that analyze the information of a set by breaking it up into its maximal consistent subsets."

So, one drawback of Lozinskii' measure is that it is a "local" one, that takes into account subsets of the whole base, but not the base as a whole for the computation of the measure of information, whereas Knight's measures are "global" ones since they keep the knowledgebase in one piece. However, the maximal consistent subset semantics seem to be natural for a lot of people (it is for example the basis for several inference relations [52, 5, 4]), and if one looks at an inconsistent knowledgebase as a knowledgebase "polluted" by some false sentences, then trying to find the "plausible" information in the maximal consistent subsets can be sensible. Finally, as underlined in section 9.1, those two measures work with sets of formulae (with a single inconsistent formula still having a null information value). This behaviour can be discussed, and can be interesting for some applications. But, in some cases, one can wish to try to get some information from an inconsistent formula. In this case, the previous measures cannot be used. In this case, the information measure proposed by Konieczny et al. (section 7), based on epistemic tests, still succeeds in extracting some non-null information from a single inconsistent formula.

As for contradiction measures, the main measures presented in this paper are, first the scoring functions of section 4. The obtained score ordering allows to compare the contradiction level of different knowledgebases. Another contradiction measure is given by Knight's $\eta-$consistency (section 6.1). The idea here is also based on minimal inconsistent subsets, and, roughly, a knowledgebase is more contradictory than another one, if the contradictions (minimal inconsistent subsets) require more formulae. Intuitively, the more formulae are needed to produce a contradiction, the less the contradiction is strong. Knight illustrates this idea on the lottery paradox: saying that, if there is a sufficiently large number of tickets, a given lottery ticket is[6] not winning, but it is a fact that one of the tickets will win the lottery. That can be written: $\Gamma = \{\neg w_1, \ldots, \neg w_n, w_1 \vee \ldots \vee w_n\}$. This knowledgebase is clearly inconsistent, but it seems sensible to say that the bigger n, the more tolerable the inconsistency. As for information measures, the two previous approaches trivialize when they are applied to singleton knowledgebase, i.e. to only one formula. They both give the maximal contradiction value to an inconsistent formula. And, even more arguably than for information measures, it is important to be able to discriminate several inconsistent formulae. This can be achieved by the degree of contradiction proposed by Konieczny et al. (section 7), that do not take the formulae as atomic inconsistencies, but take propositional variables to this aim. In this framework, it is the (maximum) number of tests required for get rid of all inconsistencies that determine the degree of contradiction. In the same way, Grant's degree of inconsistency (section 8.1) and Hunter's degree of coherence take the propositional variable as atomic inconsistency, so they can cope with singleton inconsistent knowledgebase. Those two approaches can be used to measure the amount of contradiction in a knowledgebase. But in both cases, the amount of contradiction is related to the amount of information of the base. Basically, a knowledgebase is less contradictory than another one if it

[6] more exactly it is rational to believe that a given (random) ticket is not winning.

has a model that has a lowest noise[7] ratio than the models of the second one. The main difference between the two approaches lies in the underlying chosen logic. To be able to talk about models for classically inconsistent knowledgebase, one has to choose an underlying paraconsistent logic. Grant starts from Levesque 3-valued logic [47, 26], whereas Hunter starts from quasi-classical logic [9, 29]. Quasi-classical logic seems more adequate to handle inconsistent information (it has, for example, a more constrained semantics), so the degree of coherence of Hunter may seem more adequate that Grant's degree of inconsistency. Finally let us note the degree of significance of section 8.2, that allows us to compare the amount of contradiction in several knowledgebases when the potential contradictions are not as important. It often happens in real life that some parts of the agents beliefs are more important than others[8], so contradictions that concern the important beliefs are much more problematic than the ones concerning the less important ones. So if one can weight the importance of the (worries induced by potential) contradictions, the degree of significance allows us to measure the contradiction amount of a given knowledgebase.

10 Towards Applications

Formalisation of analyses of inconsistency information has been driven by more intelligent techniques for handling inconsistent information in applications. In this section, we briefly review two emerging applications, namely negotiation between agents and comparing heterogeneous sources of information, using two of the techniques we have presented in this review.

10.1 Negotiation Between Agents

For the following example of negotiation, we will keep the domain knowledge separate from the perspectives of the participants. In other words, we will consider the domain knowledge as being correct and not subject to negotiation. This will allow us to focus our attention on the perspectives of the participants. Note, we are not presenting a general framework for negotiation between agents. Rather we are trying to show how measurement of inconsistency can be used to evaluate each cycle in a negotiation to gauge how well the negotiation is proceeding. Formalisation of multi-agent negotiation is currently the subject of much research (see for example [2, 56]). Potentially, measures of inconsistency can be incorporated in an existing formalisation for multi-agent negotiation.

Example 38. Consider three members of a family who are discussing their wishes for their next family car. Let the domain knowledge Ψ be:

[7] amount of "contradiction" compared to the amount of "information".

[8] For example, an agent can posses some beliefs that have no importance for its goals, thus having very small importance.

$$\begin{array}{l}\texttt{red} \rightarrow \texttt{fast}\\ \texttt{fast} \rightarrow \neg\texttt{fuelEfficient}\\ \texttt{offRoad} \rightarrow \texttt{expensive}\\ \texttt{sporty} \rightarrow (\texttt{expensive} \wedge (\texttt{black} \vee \texttt{red} \vee \texttt{white}))\\ \neg\texttt{expensive} \rightarrow \texttt{under\$20K}\\ \texttt{cabriolet} \rightarrow \neg\texttt{bigCapacity}\\ \texttt{fuelEfficient} \rightarrow \neg\texttt{offRoad}\end{array}$$

Let the initial preferences (requirements or demands) for each family member (participant 1, participant 2, and participant 3) be represented by Φ_1^1, Φ_1^2 and Φ_1^3 respectively.

$$\begin{array}{l}\Phi_1^1 = \{\texttt{red}, \texttt{offRoad}\}\\ \Phi_1^2 = \{\neg\texttt{expensive}, \texttt{fuelEfficient}\}\\ \Phi_1^3 = \{\texttt{sporty}, \texttt{cabriolet}, \texttt{bigCapacity}\}\end{array}$$

So the starting point of the discussions is captured by Δ_1.

$$\Delta_1 = \Psi \cup \Phi_1^1 \cup \Phi_1^2 \cup \Phi_1^3$$

Let S_1 be the scoring function for Δ_1. Now consider S_1 for some subsets of Δ_1.

$$\begin{array}{ll}S_1(\{\texttt{red}\}) = 1 & S_1(\{\texttt{bigCapacity}\}) = 1\\ S_1(\{\texttt{sporty}\}) = 1 & S_1(\{\texttt{offRoad}\}) = 2\\ S_1(\{\texttt{fuelEfficient}\}) = 2 & S_1(\{\neg\texttt{expensive}\}) = 2\\ S_1(\{\texttt{cabriolet}\}) = 1 & S_1(\{\texttt{red}, \texttt{bigCapacity}\}) = 2\end{array}$$

$$\begin{array}{l}S_1(\Phi_1^1) = S_1(\{\texttt{red}, \texttt{offRoad}\}) = 3\\ S_1(\Phi_1^2) = S_1(\{\neg\texttt{expensive}, \texttt{fuelEfficient}\}) = 4\\ S_1(\Phi_1^3) = S_1(\{\texttt{sporty}, \texttt{cabriolet}, \texttt{bigCapacity}\}) = 2\\ S_1(\Delta_1) = 5\end{array}$$

We see from S_1 that each of the preferences is individually inconsistent with the domain knowledge. We also see that Φ_1^2 has the highest score (4) of the initial preferences and it would be a good starting point for discussion.

Suppose after some discussion, Φ_1^1 is changed to Φ_2^1 by participant 1, Φ_1^2 to Φ_2^2 by participant 2, and Φ_1^3 to Φ_2^3 by participant 3, as follows. How this multi-agent discussion is conducted is beyond the scope of this review. Potential formalisms for this include [2, 56]. However, we do assume that aim of the multi-agent discussion is that some of the agents have weakened their positions. The measurement of inconsistency is intended to monitor this.

$$\begin{array}{l}\Phi_2^1 = \{\texttt{red} \vee \texttt{black}, \texttt{sporty} \vee \texttt{offRoad}\}\\ \Phi_2^2 = \{\neg\texttt{expensive}\}\\ \Phi_2^3 = \{\texttt{sporty}, \texttt{bigCapacity}\}\end{array}$$

This intermediate point is captured by Δ_2.

$$\Delta_2 = \Psi \cup \Phi_2^1 \cup \Phi_2^2 \cup \Phi_2^3$$

Let S_2 be the scoring function for Δ_2. Now consider S_2 for some subsets of Δ_2.

$$\begin{array}{l} S_2(\{\mathtt{sporty}\}) = 1 \\ S_2(\{\neg\mathtt{expensive}\}) = 2 \\ S_2(\{\mathtt{sporty} \vee \mathtt{offRoad}\}) = 1 \\ S_2(\Delta_2) = 2 \end{array}$$

We see that $S_2 < S_1$. Furthermore, we see that the preference for $\neg\mathtt{expensive}$ is the most problematical.

Now suppose after further discussion, Φ_2^2 is changed to Φ_3^2 by participant 2, and Φ_2^3 is changed to Φ_3^3 by participant 3.

$$\begin{array}{l} \Phi_3^1 = \{\mathtt{red} \vee \mathtt{black}, \mathtt{sporty} \vee \mathtt{offRoad}\} \\ \Phi_3^2 = \{\mathtt{interestFreeCredit}, \mathtt{diesel}\} \\ \Phi_3^3 = \{\mathtt{sporty} \vee \mathtt{offRoad}, \mathtt{bigCapacity}\} \end{array}$$

This final situation is captured by Δ_3.

$$\Delta_3 = \Psi \cup \Phi_3^1 \cup \Phi_3^2 \cup \Phi_3^3$$

Let S_3 be the scoring function for Δ_3. We see that $S_3 < S_2$. Also for all $\Gamma \in \Delta_3$, we have $S_3(\Gamma) = 0$. So Δ_3 could be regarded as an acceptable end-point.

In the above example, we see that the scoring functions allow us to focus on the more problematical data, and use this to facilitate conflict resolution.

10.2 Comparing Heterogeneous Sources

We now return to the problem of comparing sources, discussed in the introduction. Here we consider how the compromise relation introduced in Section 4 can be used directly to reject sources of information that are too inconsistent. A threshold can be fixed and any source of information that is above this threshold is automatically rejected. For example, if we set the threshold at 0.5, then any report represented as a set of formulae Φ that together with background knowledge Ψ is such that coherence of $\Phi \cup \Psi < 0.5$, then "more than half of the information" in Φ is contradictory with respect to the background knowledge. Similarly, for infinite models, a selected profile can be used as a threshold for rejection of sources of information.

Definition 34. *Let $\Phi_i, \Phi_j, \Psi \in \wp(\mathcal{L})$. A* **qualified compromise relation** $\preceq_\Psi$ *is defined as follows, where Φ_i and Φ_j are sources and Ψ is background knowledge.*

$$\Phi_i \preceq_\Psi \Phi_j \text{ iff } \Phi_i \cup \Psi \preceq \Phi_j \cup \Psi$$

When using a qualified compromise relation, there may be an assumption that the background knowledge is correct, and we rank sources by their conflicts with the background knowledge.

Example 39. Let Δ incorporate a standard axiomatization for the equality predicate, denoted $=$, and the "less-than-or-equal-to" predicate, denoted $\leq$. Also suppose we know that the list price of a new Ferrari Maranello is $\$200K$. We represent this as $\texttt{Cost}(\texttt{Ferrari}) = \$\texttt{200K}$, and add this to the following background knowledge in Δ.

$$\begin{array}{l} \forall \texttt{X}\ \texttt{Cost}(\texttt{X}) \leq \$\texttt{1K} \rightarrow \texttt{Cost}(\texttt{X}) \leq \$\texttt{2K} \\ \forall \texttt{X}\ \texttt{Cost}(\texttt{X}) \leq \$\texttt{2K} \rightarrow \texttt{Cost}(\texttt{X}) \leq \$\texttt{3K} \\ \qquad \vdots \\ \forall \texttt{X}\ \texttt{Cost}(\texttt{X}) \leq \$\texttt{199K} \rightarrow \texttt{Cost}(\texttt{X}) \leq \$\texttt{200K} \end{array}$$

In general, the lower the purported value of a Ferrari in a report, the greater the number of formulae in the background knowledge that are contradicted. Now consider Report 1 with the information $\texttt{Cost}(\texttt{Ferrari}) = \$\texttt{150K}$ and Report 2 with the information $\texttt{Cost}(\texttt{Ferrari}) = \$\texttt{15K}$. With this, we see Report 1 is a preferred compromise to Report 2, and that Report 1 with Δ is more coherent than Report 2 with Δ.

$$\texttt{Cost}(\texttt{Ferrari}) = \$\texttt{150K} \preceq_{\Delta} \texttt{Cost}(\texttt{Ferrari}) = \$\texttt{15K}$$

We could extend the above example so that we have the following holding for any numbers $\texttt{V}_1$ and $\texttt{V}_2$ when $\texttt{V}_1 \leq \texttt{V}_2$.

$$\texttt{Cost}(\texttt{Ferrari}) = \$\texttt{V}_1 \preceq_{\Delta} \texttt{Cost}(\texttt{Ferrari}) = \$\texttt{V}_2$$

The situation above is reflected in many real-world situations where there is a range of possible values for the facts that are being reported, and the facts that take values "further away" from those delineated by the background knowledge are regarded as more inconsistent.

As an alternative approach to dealing with heterogeneous sources, we may assume that the sources are all individually consistent with the background knowledge, but combinations of sources are inconsistent. The $\preceq$ or $\preceq_{\Psi}$ relations may then be used over all possible unions of sources. In either case, we may then choose to select the n least compromised sources of information. These n sources could then be used in some form of merging process such as arbitration [41, 42].

11 Discusssion

Current techniques for measuring the degree of inconsistency in a set of formulae are underdeveloped. There has been a marked increased interest in the past three years as reflected by new published articles on the subject. This has resulted in a range of interesting proposals based on syntactic coherence, information theory, probability theory, epistemic actions, and three/four-valued models. But it is a subject that is very much in flux. At this stage it is unclear what would constitute an ideal framework for measuring inconsistency: Though it seems that there is no unique measure of inconsistency. There are good arguments for a variety of factors to be taken into account.

Concerning the degree of inconsistency, there are two main ideas developed independently in the approaches presented in this paper. The first idea is to state that the importance of the conflict is reflected by the number of formulae of the knowledgebases implied in the contradiction. The more formulae needed, the less important the conflict. Another idea is to state that the importance of the conflict is described by the number of atoms on which we have contradictory information. An interesting question, in the quest for definition of "the" degree of consistency, is to know if it is possible to meld these two ideas, in order to take these two sensible intuitions into account.

Suggestions for desirable properties are at a tentative stage. More interrelationships between proposals need to be established. And perhaps most significantly, potential applications need to be developed. Since the more we know about how they can or should be used, the better we can develop the formalisms. In addition, there is a need to consider how some other formalisms in knowledge representation and reasoning are relevant to the subject.

Other formalisms in knowledge representation and reasoning that touch on the subject include: Diagnostic systems for which there are preferences for certain kinds of consistent subsets of inconsistent information [36, 61]; Belief revision for which epistemic entrenchment is an ordering over formulae which reflects the preference for which formulae to give up in case of inconsistency [24] and the Dalal distance which provides a model-theoretic characterisation of how inconsistent a formulae is with a consistent set of formulae [15]; Coherence-based reasoning (for drawing inferences from inconsistent information) for which there is a preference for inferences from some consistent subsets (e.g. [11, 7]); Paraconsistent logics for which there is an object operator denoting "acceptable" inconsistency that can be used to differentiate acceptable and unacceptable inconsistencies [13]; Approximate entailment for which two sequences of entailment relation are defined (the first is sound but not complete, and the second is complete but not sound) which converge to classical entailment [65]; and Partial consistency checking for which checking is terminated after the search space exceeds a threshold and so gives a measure of partial consistency of the data [54]. Whilst none of these proposals provide a direct definition for degree of inconsistency, there are clearly some important issues in common that could be explored.

As to the choice of a particular degree of inconsistency or degree of information, one important criterion, not mentionned until now, is its computational complexity. So a study of the complexity of the different proposals exposed in this paper should be a valuable work. And an open question is to know if there is a correlation between the discriminating power of the different approaches and their computational compelxity.

Finally, an interesting proposal for analysing the coherence of explanations could form an interesting development of the consistency-based analysis in Section 4. In the process of finding an explanation for some observations, there may be multiple theories that are mutually incompatible, but each constitutes an explanation for the observations. Consider a set of observation and a set of possible explanations Δ. A set $\Gamma \subseteq \Delta$ is a support for O in some context I iff $\Gamma \cup I$ implies O and no subset does so. Now we may have a number of these supports for O, and we may wish to evaluate the quality of the formulae that are used in them. In [45], a general framework for measuring support coherence is based on the average use of formulae in the supports. Highly coherent theories are those whose formulae that are tightly coupled to accounts for observations, while low coherence theories may contain disjointed and isolated statements.

Acknowledgements

The authors are grateful to Eliezer Lozinskii, Kevin M. Knight, and the anonymous referees for helpful feedback on this paper. The second author is supported by the IUT de Lens, the Université d'Artois, the Région Nord/Pas-de-Calais, and by the European Community FEDER Program. Both authors are grateful to the Royal Society and CNRS for funding travel while collaborating on this review.

References

1. C. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.
2. L. Amgoud, S. Parsons, and N. Maudet. Arguments, dialogue, and negotiation. In *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI'2000)*, 2000.
3. K. J. Arrow. *Social choice and individual values.* Wiley, New York, second edition, 1963.
4. C. Baral, S. Kraus, J. Minker, and V. S. Subrahmanian. Combining knowledge bases consisting of first-order theories. *Computational Intelligence*, 8(1):45–71, 1992.
5. C. Baral, S. Kraus, and J. Minker. Combining multiple knowledge bases. *IEEE Transactions on Knowledge and Data Engineering*, 3(2):208–220, 1991.
6. N. Belnap. A useful four-valued logic. In G Epstein, editor, *Modern Uses of Multiple-valued Logic*, pages 8–37. Reidel, 1977.
7. S. Benferhat, D. Dubois, and H. Prade. Argumentative inference in uncertain and inconsistent knowledge bases. In *Proceedings of Uncertainty in Artificial Intelligence*, pages 1449–1445. Morgan Kaufmann, 1993.
8. S. Benferhat, D. Dubois, and H. Prade. An overview of inconsistency-tolerant inferences in prioritized knowledge bases. In *Fuzzy Sets, Logic and Reasoning about Knowledge*, volume 15 of *Applied Logic Series*, pages 395–417. Kluwer, 1999.
9. Ph. Besnard and A. Hunter. Quasi-classical logic: Non-trivializable classical reasoning ¿from inconsistent information. In C Froidevaux and J Kohlas, editors, *Symbolic and Quantitative Approaches to Uncertainty*, volume 946 of *Lecture Notes in Computer Science*, pages 44–51, 1995.

10. Ph. Besnard and A. Hunter. Introduction to actual and potential contradictions. In Ph Besnard and A. Hunter, editors, *Handbook of Defeasible Resoning and Uncertainty Management Systems*, volume 2, pages 1–9. Kluwer, 1998. (Series editors: Dov Gabbay and Ph Smets).
11. G. Brewka. Preferred subtheories: An extended logical framework for default reasoning. In *Proceedings of the Eleventh International Conference on Artificial Intelligence*, pages 1043–1048, 1989.
12. L. Cholvy and C. Garion, Querying Several Conflicting Databases. *Journal Of Applied Non-Classical Logics*, 14(3):295–327, 2004.
13. W. Carnielli and J. Marcos. A taxonomy of C systems. In *Paraconsistency: The Logical Way to the Inconsistent*, pages 1–94. Marcel Dekker, 2002.
14. C. Chesnevar, A. Maguitman, and R. Loui. Logical models of argument. *ACM Computing Surveys*, 32:337–383, 2001.
15. M. Dalal. Investigations into a theory of knowledge base revision: Preliminary report. In *Proceedings of the 7th National Conference on Artificial Intelligence (AAAI'88)*, pages 3–7. MIT Press, 1988.
16. A. Darwiche and J. Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89:1–29, 1997.
17. N. C. da Costa. On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic*, 15:497–510, 1974.
18. D. Dubois, S. Konieczny, and H. Prade. Quasi-possibilistic logic and its measures of information and conflict. *Fundamenta Informaticae*, 57:101–125, 2003.
19. D. Dubois, J. Lang, and H. Prade. Possibilistic logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3, pages 439–513. Oxford University Press, 1994.
20. D. Dubois and H. Prade. Properties of measures of information in evidence and possibility theories. *Fuzzy Sets and Systems*, 24:161–182, 1987. Reprinted in Fuzzy Sets and Systems, supplement to Vol. 100, 35-49, 1999.
21. M. Elvang-Goransson and A. Hunter. Argumentative logics: Reasoning from classically inconsistent information. *Data and Knowledge Engineering*, 16:125–145, 1995.
22. N. Friedman and J. Y. Halpern. Belief revision: a critique. In *Proceedings of the Fifth International Conference on Principles of Knowledge Representation and Reasoning (KR'96)*, pages 421–431, 1996.
23. D. Gabbay and A. Hunter. Making inconsistency respectable 1: A logical framework for inconsistency in reasoning. In *Fundamentals of Artificial Intelligence*, volume 535 of *Lecture Notes in Computer Science*, pages 19–32. Springer, 1991.
24. P. Gärdenfors. *Knowledge in Flux*. MIT Press, 1988.
25. J. Grant and A. Hunter. Measuring inconsistency in knowledgebases. Technical report, UCL Department of Computer Science, 2004.
26. J. Grant. Classifications for inconsistent theories. *Notre Dame Journal of Formal Logic*, 19:435–444, 1978.
27. J. Hintikka. On semantic information. *Information and Inference*, pages 3–27, 1970.
28. A. Hunter. Paraconsistent logics. In Dov Gabbay and Ph Smets, editors, *Handbook of Defeasible Resoning and Uncertainty Management Systems*, volume 2, pages 11–36. Kluwer, 1998.
29. A. Hunter. Reasoning with contradictory information using quasi-classical logic. *Journal of Logic and Computation*, 10:677–703, 2000.

30. A. Hunter. A semantic tableau version of first-order quasi-classical logic. In *Quantitative and Qualitative Approaches to Reasoning with Uncertainty*, LNCS. Springer, 2001. 544–556.
31. A. Hunter. Measuring inconsistency in knowledge via quasi-classical models. In *Proceedings of the 18th National Conference on Artificial Intelligence (AAAI'2002)*, pages 68–73. MIT Press, 2002. ISBN 0-262-51129-0.
32. A. Hunter. Evaluating the significance of inconsistency. In *Proceedings of the International Joint Conference on AI (IJCAI'03)*, pages 468–473, 2003.
33. A. Hunter. Measuring inconsistency in first-order knowledge. Technical report, UCL Department of Computer Science, 2003.
34. A. Hunter. Logical comparison of inconsistent perspectives using scoring functions. *Knowledge and Information Systems Journal*, 2004. (in press).
35. J. Kemeny. A logical measure function. *Journal of Symbolic Logic*, 18:289–308, 1953.
36. J. De Kleer and B. Williams. Diagnosing multiple faults. *Artificial Intelligence*, 32:97–130, 1987.
37. K. M. Knight. Measuring inconsistency. *Journal of Philosophical Logic*, 31:77–98, 2001.
38. K. M. Knight. *A theory of inconsistency*. PhD thesis, The university of Manchester, 2002.
39. K. M. Knight. Two information measures for inconsistent sets. *Journal of Logic, Language and Information*, 12:227–248, 2003.
40. S. Konieczny, J. Lang, and P. Marquis. Quantifying information and contradiction in propositional logic through epistemic tests. In *Proceedings of the 18th International Joint Conference on Artificial Intellignce (IJCAI'03)*, pages 106–111, 2003.
41. S. Konieczny and R. Pino-Pérez. On the logic of merging. In *Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning (KR98)*, pages 488–498. Morgan Kaufmann, 1998.
42. S. Konieczny and R. Pino-Pérez. Merging information under constraints: a qualitative framework. *Journal of Logic and Computation*, 12(5):773–808, 2002.
43. S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.
44. T. Kuipers, editor. *What is closer-to-the-truth?* Rodopi, Amsterdam, 1987.
45. R. Kwok, N. Foo, and A. Nayak. Coherence of laws. In *Proceedings of the International Joint Conference on AI (IJCAI'03)*, 2003.
46. D. Lehmann and M. Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55:1–60, 1992.
47. H. Levesque. A logic of implicit and explicit belief. In *Proceedings of the National Conference on Artificial Intelligence (AAAI'84)*, pages 198–202, 1984.
48. E. Lozinskii. Information and evidence in logic systems. *Journal of Experimental and Theoretical Artificial Intelligence*, 6:163–193, 1994.
49. E. Lozinskii. Resolving contradictions: A plausible semantics for inconsistent systems. *Journal of Automated Reasoning*, 12:1–31, 1994.
50. E. Lozinskii. On knowledge evolution: Acquisition, revision, contraction. *Journal of Applied Non-Classical Logic*, 7:177–212, 1997.
51. D. Makinson. *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume III, chapter General Pattern in nonmonotonic reasoning, pages 35–110. Clarendon Press, Oxford, 1994.
52. R. Manor and N. Rescher. On inferences from inconsistent information. *Theory and Decision*, 1:179–219, 1970.

53. P. Marquis and N. Porquet. Computational aspects of quasi-classical entailment. *Journal of Applied Non-classical Logics*, 11:295–312, 2001.
54. C. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.
55. J. B. Paris. *The uncertain reasoner's companion: a mathematical perspective*. Number 39 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1994.
56. S. Parsons, M. Wooldridge, and L. Amgoud. Properties and complexity of some formal inter-agent dialogues. *Journal of Logic and Computation*, 2003.
57. K. Popper. *Conjectures and Refutation*. Routlege and Kegan Paul, London, 1963.
58. H. Prakken and G. Vreeswijk. Logical systems for defeasible argumentation. In D Gabbay, editor, *Handbook of Philosophical Logic*. Kluwer, 2000.
59. G. Priest. Minimally inconsistent LP. *Studia Logica*, 50:321–331, 1991.
60. G. Priest. Paraconsistent logic. In *Handbook of Philosophical Logic*, volume 6. Kluwer, 2002.
61. R. Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32:57–95, 1987.
62. P. Revesz. On the semantics of arbitration. *International Journal of Algebra and Computation*, 7:133–160, 1997.
63. M. Ryan and P. Y. Schobbens. Belief revision and verisimilitude. *Notre Dame Journal of Formal Logic*, 36(1):15–29, 1995.
64. L. J. Savage. *The foundations of statistics*. Dover Publications, New York, 1971. Second revised edition.
65. M. Schaerf and M. Cadoli. Tractable reasoning via approximation. *Artificial Intelligence*, 74:249–310, 1995.
66. G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
67. C. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.
68. W. Spohn. Ordinal conditional functions: a dynamic theory of epistemic states. In W. L. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change, and Statistics*, volume 2, pages 105–134. Princeton University Press, 1987.
69. M. Williams. Iterated theory base change: A computational model. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI'95)*, pages 1541–1547, 1995.
70. P. Wong and Ph. Besnard. Paraconsistent reasoning as an analytic tool. *Journal of the Interest Group in Propositional Logic*, 9:233–246, 2001.