Measuring Inconsistency in Knowledge via Quasi-classical Models

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

Abstract

The language for describing inconsistency is underdeveloped. If a knowledgebase (a set of formulae) is inconsistent, we need more illuminating ways to say how inconsistent it is, or to say whether one knowledgebase is "more inconsistent" than another. To address this, we provide a general characterization of inconsistency, based on quasi-classical logic (a form of paraconsistent logic with a more expressive semantics than Belnap's four-valued logic, and unlike other paraconsistent logics, allows the connectives to appear to behave as classical connectives). We analyse inconsistent knowledge by considering the conflicts arising in the minimal quasi-classical models for that knowledge. This is used for a measure of coherence for each knowledgebase, and for a preference ordering, called the compromise relation, over knowledgebases. In this paper, we formalize this framework, and consider applications in managing heterogeneous sources of knowledge.

Introduction

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.

Current techniques for measuring the degree of inconsistency in a set of formulae are underdeveloped. Some approaches touch on the topic. In diagnostic systems, there are proposals that offer preferences for certain kinds of consistent subsets of inconsistent information (Kleer & Williams 1987; Reiter 1987); in proposals for belief revision, epistemic entrenchment is an ordering over formulae which reflects the preference for which formulae to give up in case of inconsistency (Gardenfors 1988); in proposals for drawing inferences from inconsistent information there is a preference for inferences from some consistent subsets (e.g. (Brewka 1989; Benferhat, Dubois, & Prade 1993)); in proposals for approximating entailment, 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 (Schaerf & Cadoli 1995); and in proposals for partial consistency checking, checking is terminated after the search space exceeds a threshold which gives a measure of partial consistency of the data. However, none of these proposals provide a direct definition for degree of inconsistency.

In belief revision theory, and the related field of knowledgebase merging, there are some proposals that do provide some description of the degree of inconsistency of a set of formulae. For example, the Dalal distance (Dalal 1988), essentially the Hamming distance between two propositional interpretations, can be used to give a profile of an inconsistent knowledgebase. Let $dalal(w, w')$ denote the Dalal distance from w to w', let $[\alpha]$ denote the set of classical models of α, and let $d(w, \alpha)$ be the $w' \in [\alpha]$ such that $dalal(w, w')$ is minimized. Now suppose we have a knowledgebase $\{\alpha_1, .., \alpha_n\}$ where each α_i is consistent but the knowledgebase may be inconsistent. We can then obtain a value of $d(w, \alpha_i)$ for each world w and each formula α_i in the knowledgebase. Unfortunately, this does not provide a very succinct way of describing the degree of inconsistency in a given set of formulae, and it is not clear how we could compare sets of formulae using this approach. Furthermore, operators for aggregating these distances, such as the majority operator (Lin & Mendelzon 1998), egalitarist operator (Revesz 1997), or the leximax operator (Konieczny & Pino Perez 1998), do not seem to be appropriate summaries of the degree of inconsistency in the original knowledgebase since they seek to find the most appropriate model for particular kinds of compromise of the original knowledge. Related techniques for knowledgebase revision are similarly inappropriate.

Another approach to handling inconsistent information is that of possibility theory (Dubois, Lang, & Prade 1994). 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. As presented, the problem with this measure is that it assumes weighted formulae. In other words, we need some form of preference ordering in addition to the set of classical formulae in the knowledgebase. The knowledgebase can be used to induce such an ordering as suggested in (Benferhat *et al.* 2000), where an ordering over inferentially weaker forms of the original formulae are generated. Again this does not offer a direct lucid view on the inconsistency in the original set of formulae.

Measuring the "amount of information" is related to the idea of measuring inconsistency. Information theory can be used to measure the information content of sets of inconsistent formulae. Applying Shannon's measure of information, Lozinskii proposes that the information in a set of propositional formulae Γ, that has been 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 (Lozinskii 1994). This information theoretic measure increases with additions of consistent information and decreases with additions of inconsistent information. However, as highlighted by Wong and Besnard, the measure by Lozinskii is syntax sensitive and it is sensitive to the presence of tautologies in Γ. To address this, they suggest the use of a normal form for the formulae in Γ that is obtained by rewriting Γ into conjunctive normal form, and then applying disjunction elimination and resolution exhaustively (Wong & Besnard 2001). However, this approach does not provide a direct measure of inconsistency since for example, the value for $\{\alpha\}$ is the same as for $\{\alpha, \neg\alpha, \beta\}$.

In this paper, we want to reflect each inconsistent set of formulae in a model, and then measure the inconsistency in the model. Obviously, this is not possible in classical logic, or indeed many non-classical logics, because there is no model of an inconsistent set of formulae. We therefore turn to quasi-classical logic, a form of paraconsistent logic, to model inconsistent sets of formulae. There are other paraconsistent logics that we could consider, for example Belnap's four-valued logic (Belnap 1977), or Levesque's 3-interpretations (Levesque 1984), or Grant's generalizations of classical satisfaction (Grant 1978), but these, as we will illustrate, involve the consideration of too many models. This increases the number of models that need to be analysed and it underspecifies the nature of the conflicts.

In this paper, we review the aspects of QC logic that we will require for the rest of the paper, we argue why QC models are more appropriate than those obtained from other paraconsistent logics, we define a new framework for measuring inconsistencies in models, and we extend this semantic framework to preference relations over sets of formulae.

Review of QC Logic

We review the propositional version of quasi-classical logic (QC Logic) (Besnard & Hunter 1995; Hunter 2000).

Definition 1 *The* **language of first-order QC logic** *is that of classical propositional logic. We let $\mathcal{L}$ denote a set of formulae formed in the usual way from a set of atom symbols $\mathcal{A}$, and the connectives $\{\neg, \vee, \wedge, \rightarrow\}$. If $\Gamma \in \wp(\mathcal{L})$, then* $\mathsf{Atoms}(\Gamma)$ *returns the set of atom symbols used in Γ.*

Definition 2 *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 3 *Let $\alpha_1 \vee .. \vee \alpha_n$ be a clause that includes a literal 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 1 *Let $\alpha \vee \beta \vee \gamma$ be a clause where α, β, and γ are literals. Hence, $\otimes(\alpha \vee \beta \vee \gamma, \beta) = \alpha \vee \gamma$.*

We now consider the essential idea behind QC logic. We describe it using the resolution proof rule. Resolution can be applied to clauses to generate further clauses called resolvents. For example, by resolution $\beta \vee \gamma$ is a resolvent of $\alpha \vee \beta$ and $\neg\alpha \vee \gamma$. Given a set of clauses as assumptions, each clause in the assumptions can be regarded as a belief, and each resolvent can be regarded as a belief. So resolution can be regarded as a process of focusing beliefs.

A useful property of resolution is that α is a resolvent only if all the literals used in α are literals used in the set of assumptions (assuming no introduction proof rules are used). This means that any resolvent, and hence any belief derivable from the assumptions, is a non-trivial inference from the assumptions. This holds even if the set of assumptions is classically inconsistent. As a result, resolution can constitute the basis of useful paraconsistent reasoning.

QC logic is motivated by the need to handle beliefs rather than the need to address issues of verisimilitude for given propositions. It is intended to be a logic of beliefs in the "real world" rather than a logic of truths in the "real world". Models are based on a form of Herbrand interpretation.

Definition 4 *Let $\mathcal{A}$ be a set of atoms. Let $\mathcal{O}$ be the set of objects defined as follows, where $+\alpha$ is a positive object, and $-\alpha$ is a negative object.*

$$\mathcal{O} = \{+\alpha \mid \alpha \in \mathcal{A}\} \cup \{-\alpha \mid \alpha \in \mathcal{A}\}$$

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 that in X there is **a reason against** the belief $\neg\alpha$. Similarly, $-\alpha \in X$ means that in X there is **a reason against** the belief α and that in X there is **a reason for** the belief $\neg\alpha$.

Definition 5 *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 $\mathcal{L}$.*

$X \models_s \alpha$ *iff there is a reason for the belief α in X*

$X \models_s \alpha_1 \vee ... \vee \alpha_n$
iff $[X \models_s \alpha_1$ *or ... or* $X \models_s \alpha_n]$
and $\forall i$ *s.t.* $1 \leq i \leq n$
$[X \models_s \sim\alpha_i$ *implies* $X \models_s \otimes(\alpha_1 \vee ... \vee \alpha_n, \alpha_i)]$

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

$X \models_s \alpha \wedge \beta$ *iff* $X \models_s \alpha$ *and* $X \models_s \beta$
$X \models_s \neg\neg\alpha \vee \gamma$ *iff* $X \models_s \alpha \vee \gamma$
$X \models_s \neg(\alpha \wedge \beta) \vee \gamma$ *iff* $X \models_s \neg\alpha \vee \neg\beta \vee \gamma$
$X \models_s \neg(\alpha \vee \beta) \vee \gamma$ *iff* $X \models_s (\neg\alpha \wedge \neg\beta) \vee \gamma$
$X \models_s \alpha \vee (\beta \wedge \gamma)$ *iff* $X \models_s (\alpha \vee \beta) \wedge (\alpha \vee \gamma)$
$X \models_s \alpha \wedge (\beta \vee \gamma)$ *iff* $X \models_s (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$
$X \models_s (\alpha \rightarrow \beta) \vee \gamma$ *iff* $X \models_s \neg\alpha \vee \beta \vee \gamma$
$X \models_s \neg(\alpha \rightarrow \beta) \vee \gamma$ *iff* $X \models_s (\alpha \wedge \neg\beta) \vee \gamma$

Definition 6 *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 2 *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.*

The following result from (Hunter 2000) provides a slightly different view on the semantics of disjunction.

Proposition 1 *Let $X \in \wp(\mathcal{O})$, and $\alpha_1, .., \alpha_n$ be literals in $\mathcal{L}$. We have $X \models_s \alpha_1 \vee .. \vee \alpha_n$ iff (1) for some $\alpha_i \in \{\alpha_1, .., \alpha_n\}$, $X \models_s \alpha_i$ and $X \not\models_s \sim\alpha_i$ or (2) for all $\alpha_i \in \{\alpha_1, .., \alpha_n\}$, $X \models_s \alpha_i$ and $X \models_s \sim\alpha_i$.*

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 (Hunter 2000) and a semantic tableau version for first-order QC logic (Hunter 2001). 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 (Marquis & Porquet 2001).

Why measure inconsistency with QC models?

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 (da Costa 1974), and consistency-based logics such as (Benferhat, Dubois, & Prade 1993). Furthermore, the semantics of QC logic directly models inconsistent sets of formulae.

Whilst four-valued logic (Belnap 1977) also directly models inconsistent sets of formulae, QC logic is stronger in the sense that the number of non-tautological inferences obtained from a set of formulae is never less than with four-valued logic, and often it is greater. Consider the example $\{\alpha \vee \beta, \neg\alpha\}$ from which the inference β can be obtained with QC logic but not with four-valued logic. This stronger notion of inference and entailment is reflected in the models and so more closely reflects the non-trivial aspects of classical reasoning.

Another way of viewing the weakness of Belnap's four-valued logic is that there are too many four-valued models in many situations. Consider for example $\{\alpha \vee \beta, \neg\alpha\}$. We have one minimal QC model $\{-\alpha, +\beta\}$, but with four-valued logic there are a number of 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.

Another approach that we should consider here is that of 3-interpretation by (Levesque 1984), and a similar proposal by (Grant 1978). 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. As with Belnap's four-valued logic, there are too many models. First consider $\{\alpha, \neg\alpha \vee \neg\beta, \beta\}$. This has three 3-interpretations: (1) $\alpha, \neg\alpha, \beta$ are true and $\neg\beta$ is false; (2) $\alpha, \beta, \neg\beta$ are true and $\neg\alpha$ is false; and (3) $\alpha, \neg\alpha, \beta, \neg\beta$ are true. In contrast, there is just one minimal QC model $\{+\alpha, -\alpha, +\beta, -\beta\}$, and we argue that this QC model better describes the conflicts in the set of formulae. Now consider $\{\alpha, \neg\alpha \vee \neg\beta\}$. This has three 3-interpretations: (1) $\alpha, \neg\alpha, \neg\beta$ are true; (2) $\alpha, \neg\beta$ are true and $\neg\alpha$ is false; and (3) $\alpha, \neg\alpha$ are true and $\neg\beta$ is false. In contrast, there is just one minimal QC model $\{+\alpha, -\beta\}$.

Minimal QC models

For measuring inconsistency, we use minimal QC models.

Definition 7 *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 3 *Consider the following sets of formulae.*

$$\mathsf{MQC}(\{\alpha \wedge \neg\alpha, \alpha \vee \beta, \neg\alpha \vee \gamma\}) = \{\{+\alpha, -\alpha, +\beta, +\gamma\}\}$$
$$\mathsf{MQC}(\{\neg\alpha \wedge \alpha, \beta \vee \gamma\}) = \{\{+\alpha, -\alpha, +\beta\}, \{+\alpha, -\alpha, +\gamma\}\}$$
$$\mathsf{MQC}(\{\alpha \vee \beta, \neg\alpha \vee \gamma\}) = \{\{+\beta, +\gamma\}, \{+\alpha, +\gamma\}, \{-\alpha, +\beta\}\}$$

Proposition 2 *Let $\mathsf{Atoms}(\Delta) = n$. If $X \in \mathsf{MQC}(\Delta)$, then $|X| \leq 2n$. Also, if $|X| = 2n$, and $X \in \mathsf{MQC}(\Delta)$, then $\mathsf{MQC}(\Delta) = \{X\}$.*

Increasing the number of inconsistencies in a knowledgebase tends to decrease the number of minimal QC models.

Definition 8 *Let $\Delta \in \wp(\mathcal{L})$, and $\mathsf{Incon}(\Delta) = \{\Gamma \subseteq \Delta \mid \Gamma \vdash \bot\}$ where $\vdash$ is classical consequence. The set of minimal inconsistent subsets of Δ, denoted $\mathsf{MI}(\Delta)$, is defined as,*

$$\mathsf{MI}(\Delta) = \{\Phi \in \mathsf{Incon}(\Delta) \mid \forall\Psi \in \mathsf{Incon}(\Delta)\ \Psi \not\subset \Phi\}$$

The following shows a simple relationship between the minimal QC models for a knowledgebase and those for the minimal inconsistent subsets of it.

Proposition 3 *Let* $\Delta \in \wp(\mathcal{L})$. *Let* $\Gamma \in \mathsf{MI}(\Delta)$. *If* $X \in \mathsf{MQC}(\Delta)$, *then* $\exists Y \in \mathsf{MQC}(\Gamma)$ *such that* $Y \subseteq X$. *However if* $Y \in \mathsf{MQC}(\Gamma)$, *then it is not necessarily the case that* $\exists X \in \mathsf{MQC}(\Delta)$ *such that* $Y \subseteq X$.

Finding minimal QC models is more expensive than finding just QC models. For any set of formulae Δ, let $X = \{+\alpha \mid \alpha \in \mathsf{Atoms}(\Delta)\} \cup \{-\alpha \mid \alpha \in \mathsf{Atoms}(\Delta)\}$. So X is a QC model that satisfies Δ, and it can be found in time that is a linear function of the size of the formulae, though X is not necessarily a minimal QC model. If $|X| = n$, then there are 2^n QC models that can be formed from X by taking subsets of X. This includes all minimal QC models of Δ.

Proposition 4 *Determining whether* $X \in \mathsf{MQC}(\Delta)$ *holds is a coNP-complete problem when all* $\alpha \in \Delta$ *are CNF.*

This result is based on the linear time transformation by (Marquis & Porquet 2001) that can be used to turn a QC satisfaction problem into a classical satisfaction problem, and on the coNP-complete property of checking whether a given classical interpretation is a minimal model of a given formula (Cadoli 1992).

Measuring coherence of QC models

We now consider a measure of inconsistency called coherence. 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 9 *Let* $X \in \wp(\mathcal{O})$.

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

If $\mathsf{Opinionbase}(X) = \emptyset$, then X has no arguments for/against any beliefs, and hence X has no opinions. If $\mathsf{Opinionbase}(X) = \mathcal{A}$, then X is totally opinionated. If $\mathsf{Conflictbase}(X) = \emptyset$, then X is a conflictfree QC model. If $\mathsf{Opinionbase}(X) = \mathcal{A}$, and $\mathsf{Conflictbase}(X) = \emptyset$, then we describe X as omniscient.

In finding the minimal QC models for a set of formulae, minimization of the size of each model forces minimization of the conflictbase of each model. As a result of this minimization, there is a unique conflictbase that is common to all the minimal QC models for each set of formulae, though there is not necessarily a unique opinionbase.

Proposition 5 *Let* $\Delta \in \wp(\mathcal{L})$. *If* $X, Y \in \mathsf{MQC}(\Delta)$, *then (1)* $\mathsf{Conflictbase}(X) = \mathsf{Conflictbase}(Y)$ *and (2) either* $\mathsf{Opinionbase}(X) = \mathsf{Opinionbase}(Y)$ *or* $\mathsf{Opinionbase}(X)$ *is not a subset of* $\mathsf{Opinionbase}(Y)$.

This result is based on Proposition 1, and on the following observation, where Δ is a set of clauses: For all $X \in \mathsf{MQC}(\Delta)$, $+\alpha \in X$ and $-\alpha \in X$ iff there is a $\phi \in \cup\mathsf{MI}(\Delta)$ such that α is a disjunct in ϕ and there is a $\psi \in \cup\mathsf{MI}(\Delta)$ such that $\neg\alpha$ is a disjunct in ψ.

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

Definition 10 *The* Coherence *function from* $\wp(\mathcal{O})$ *into* $[0, 1]$, *is defined 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 4 *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* $Z \in \mathsf{MQC}(\{\neg\alpha, \beta, \neg\gamma \wedge \gamma\})$. *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.

Example 5 *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 11 *Let* $\Delta \in \wp(\mathcal{L})$. *Assign* $\mathsf{Coherence}(\Delta)$ *the maximum value in* $\{\mathsf{Coherence}(X) \mid X \in \mathsf{MQC}(\Delta)\}$

Example 6 *Let* $\Delta = \{\alpha \wedge \neg\alpha, \beta \wedge \neg\beta, \alpha \vee \beta \vee (\gamma \wedge \delta)\}$ *and* $\Delta' = \{\alpha \wedge \neg\alpha, \alpha \vee \beta\}$. *Here* $\mathsf{Coherence}(\Delta) = \mathsf{Coherence}(\Delta') = 1/2$.

Example 7 *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* $\mathsf{Coherence}(\Delta) > \mathsf{Coherence}(\Delta')$.

The coherence function is not a monotonic function, as illustrated by the following example.

Example 8 *Let* $\Delta = \{\alpha\}$ *and* $\Delta' = \{\alpha, \neg\alpha, \beta\}$. *So* $\Delta \subset \Delta'$, *and* $\mathsf{Coherence}(\Delta) > \mathsf{Coherence}(\Delta')$. *Now let* $\Delta'' = \{\alpha, \neg\alpha\}$. *So* $\Delta \subset \Delta'' \subset \Delta'$, *and* $\mathsf{Coherence}(\Delta'') < \mathsf{Coherence}(\Delta')$.

The coherence function does not discriminate on the number or intersection of the minimal inconsistent subsets of a knowledgebase as illustrated by the following example.

Example 9 *Let* $\Delta = \{\alpha \wedge \neg\alpha, \beta \wedge \neg\beta\}$ *and* $\Delta' = \{\alpha \wedge \beta, \neg\alpha \wedge \neg\beta\}$. *Here* Δ *has two disjoint minimal inconsistent subsets whereas* Δ' *has one. Yet* $\mathsf{Coherence}(\Delta) = \mathsf{Coherence}(\Delta') = 0$.

Example 10 *Let* $\Delta = \{\alpha \wedge \neg\alpha\}$ *and* $\Delta' = \{\beta \wedge \neg\beta\}$. *Here* $\mathsf{Coherence}(\Delta) = \mathsf{Coherence}(\Delta')$ *even though* Δ *and* Δ' *are quite distinct as indicated by* $\mathsf{Atoms}(\Delta) \cap \mathsf{Atoms}(\Delta') = \emptyset$.

In the next section, we present an alternative to the coherence function for comparing knowledgebases. We aim to differentiate between knowledgebases such as Δ and Δ' given in Example 10.

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

$$\Delta \preceq \Delta' \text{ iff } \forall X \in \mathsf{MQC}(\Delta) \text{ and } \exists Y \in \mathsf{MQC}(\Delta') \text{ such that } \mathsf{Conflictbase}(X) \subseteq \mathsf{Conflictbase}(Y) \text{ and } \mathsf{Opinionbase}(Y) \subseteq \mathsf{Opinionbase}(X)$$

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

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

Example 12 *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 13 *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'. We get this via Proposition 6. 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 11. And according to Proposition 7, this is sufficient to ensure that there is no minimal QC model of Δ' that has a greater opinionbase than any minimal QC model of Δ.

Proposition 6 *If $\Delta \preceq \Delta'$, then $\forall X \in \mathsf{MQC}(\Delta)$, $\forall Y \in \mathsf{MQC}(\Delta')$, $\mathsf{Conflictbase}(X) \subseteq \mathsf{Conflictbase}(Y)$.*

Proposition 7 *If $\Delta \preceq \Delta'$, then it is not the case that $\exists X \in \mathsf{MQC}(\Delta)$, $\exists Y \in \mathsf{MQC}(\Delta')$, such that $\mathsf{Opinionbase}(X) \subset \mathsf{Opinionbase}(Y)$.*

Useful properties of the compromise relation include: (1) It is a pre-order relation; (2) It captures aspects of coherence (Propositions 8 and 9); and (3) It is syntax independent (Proposition 10).

Proposition 8 *If $\Delta \preceq \Delta'$, then $\forall X \in \mathsf{MQC}(\Delta)$ $\exists Y \in \mathsf{MQC}(\Delta')$ such that $\mathsf{Coherence}(X) \geq \mathsf{Coherence}(Y)$.*

Proposition 9 *If $\Delta \preceq \Delta'$, and $\Delta = \cup\mathsf{MI}(\Delta)$, and $\Delta' = \cup\mathsf{MI}(\Delta')$, and $\Delta \cup \Delta'$ is a set of clauses, then $\mathsf{Coherence}(\Delta) \geq \mathsf{Coherence}(\Delta')$.*

However, in general $\Delta \preceq \Delta'$ does not imply $\mathsf{Coherence}(\Delta) \geq \mathsf{Coherence}(\Delta')$. The converse does not hold either. This is illustrated by the following examples.

Example 14 *Let $\Delta = \{\alpha, \delta \wedge \neg\delta\}$ and $\Delta' = \{\alpha \vee (\beta \wedge \gamma), \delta \wedge \neg\delta\}$. So $\Delta \preceq \Delta'$ and $\mathsf{Coherence}(\Delta) < \mathsf{Coherence}(\Delta')$.*

Example 15 *Let $\Delta = \{\alpha, \beta \wedge \neg\beta\}$ and $\Delta' = \{\alpha, \gamma \wedge \neg\gamma\}$. So $\mathsf{Coherence}(\Delta) \geq \mathsf{Coherence}(\Delta')$. However, $\Delta \not\preceq \Delta'$ and $\Delta' \not\preceq \Delta$.*

We see that increasing the number of conjuncts in a formula tends to increase the size of the minimal QC models for that formula, whereas increasing the number of disjuncts in a formula tends to increase the number of the minimal QC models for that formula. We illustrate this in the following example, and see the effect on the compromise relation.

Example 16 *If $\Delta = \{\alpha \wedge \beta\}$ and $\Delta' = \{\alpha\}$ and $\Delta'' = \{\alpha \vee \beta\}$, then $\Delta \prec \Delta'$, and $\Delta' \prec \Delta''$, since $\mathsf{MQC}(\Delta) = \{\{+\alpha, +\beta\}\}$, $\mathsf{MQC}(\Delta') = \{\{+\alpha\}\}$, and $\mathsf{MQC}(\Delta'') = \{\{+\alpha\}, \{+\beta\}\}$.*

If Δ and Δ' are consistent sets of formulae, and $\Delta \simeq \Delta'$, then Δ and Δ' are not necessarily classically equivalent, as illustrated by the following example.

Example 17 *Let $\Delta = \{\neg\alpha\}$ and $\Delta' = \{\alpha\}$. So $\Delta \not\vdash \bot$, and $\Delta' \not\vdash \bot$ and $\Delta \simeq \Delta'$.*

The behaviour of the compromise relation illustrated above is the result of the opinionbase comparison within the compromise relation not differentiating between formulae and their complements. However, the compromise relation is syntax independent, which we formalize as follows.

Definition 13 *For $\Delta, \Delta' \in \wp(\mathcal{L})$, Δ is semantically equivalent to Δ' iff $\forall X \in \wp(\mathcal{O})(X \models_s \Delta$ iff $X \models_s \Delta')$. Let $\mathsf{SemanticEqual}$ be a function that gives the set of semantically equivalent knowledgebases for a knowledgebase.*

Example 18 *Let $\Delta = \{\alpha, \neg\alpha\}$. So $\{\alpha \wedge \neg\alpha\} \in \mathsf{SemanticEqual}(\Delta)$. Now let $\Delta' = \{\alpha, \neg\alpha \vee \neg\beta, \beta\}$. So $\{\alpha \wedge \neg\alpha, \beta \wedge \neg\beta\} \in \mathsf{SemanticEqual}(\Delta')$ and $\{\alpha, \neg\alpha \vee (\beta \wedge \neg\beta)\} \in \mathsf{SemanticEqual}(\Delta')$.*

Proposition 10 *Let $\Delta \in \wp(\mathcal{L})$. If $\Delta'' \in \mathsf{SemanticEqual}(\Delta)$, then $\Delta \simeq \Delta''$.*

Hence, $\preceq$ is syntax independent. As a result, if $\Delta \preceq \Delta'$ and $\Delta'' \in \mathsf{SemanticEqual}(\Delta)$, then $\Delta'' \preceq \Delta'$. And if $\Delta \preceq \Delta'$ and $\Delta'' \in \mathsf{SemanticEqual}(\Delta')$, then $\Delta \preceq \Delta''$.

Example 19 *Since $\{\alpha, \neg\alpha, \beta\} \preceq \{\alpha, \neg\alpha, \beta, \neg\beta\}$ holds, we can derive that $\{\alpha \wedge \neg\alpha \wedge \beta\} \preceq \{\alpha, \neg\alpha, \beta, \neg\beta\}$ holds.*

However, this syntax independence means that the compromise relation does not reflect the membership and cardinalities of the inconsistent subsets as illustrated by the following examples.

Example 20 *Let* $\Delta = \{\beta\}$ *and* $\Delta' = \{\alpha, \neg\alpha, \beta\}$. *So* $\mathsf{MI}(\Delta) = \emptyset$ *and* $\mathsf{MI}(\Delta') = \{\{\alpha, \neg\alpha\}\}$, *but* $\Delta \not\preceq \Delta'$ *and* $\Delta' \not\preceq \Delta$.

Example 21 *Let* $\Delta = \{\alpha, \neg\alpha \vee \neg\beta, \beta\}$ *and* $\Delta' = \{\neg\alpha, \alpha \vee \beta, \neg\beta\}$. *So* $\mathsf{MI}(\Delta) = \{\Delta\}$, *and* $\mathsf{MI}(\Delta') = \{\Delta'\}$, *and* $\Delta \simeq \Delta'$, *but* $\mathsf{MI}(\Delta) \not\subseteq \mathsf{MI}(\Delta')$ *and* $\mathsf{MI}(\Delta') \not\subseteq \mathsf{MI}(\Delta)$.

In general, the compromise relation is not monotonic so for instance, $\Delta \preceq \Delta'$ does not imply (1) $\Delta \cup \Gamma \preceq \Delta'$, (2) $\Delta \preceq \Delta' \cup \Gamma$, or (3) $\Delta \cup \Gamma \preceq \Delta' \cup \Gamma$. This is illustrated by the following example.

Example 22 *(1) Let* $\Delta_1 = \{\alpha, \beta\}$, $\Delta'_1 = \{\alpha\}$, *and* $\Gamma_1 = \{\neg\alpha\}$. *So* $\Delta_1 \preceq \Delta'_1$ *but* $\Delta_1 \cup \Gamma_1 \not\preceq \Delta'_1$. *(2) Let* $\Delta_2 = \{\alpha\}$, $\Delta'_2 = \{\alpha\}$, *and* $\Gamma_2 = \{\beta\}$. *So* $\Delta_2 \preceq \Delta'_2$ *but* $\Delta_2 \not\preceq \Delta'_2 \cup \Gamma_2$. *(3) Let* $\Delta_3 = \{\alpha \wedge \beta\}$, $\Delta'_3 = \{\alpha \vee \beta\}$, *and* $\Gamma_3 = \{\neg\alpha\}$. *So* $\Delta_3 \preceq \Delta'_3$ *but* $\Delta_3 \cup \Gamma_3 \not\preceq \Delta'_3 \cup \Gamma_3$.

However, if have an update Γ, where $\mathsf{Atoms}(\Delta) \cap \mathsf{Atoms}(\Gamma) = \emptyset$ and $\mathsf{Atoms}(\Delta') \cap \mathsf{Atoms}(\Gamma) = \emptyset$, then $\Delta \preceq \Delta'$ implies $\Delta \cup \Gamma \preceq \Delta' \cup \Gamma$.

Analysing heterogeneous sources

Returning to the problem of comparing sources, discussed in the introduction, we briefly consider two types of analysis.

Definition 14 *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.*

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

Example 23 *Let* $\Phi_1 = \{\neg\alpha, \neg\beta, \neg\gamma \vee \delta\}$, $\Phi_2 = \{\neg\alpha, \neg\beta, \delta, \neg\gamma\}$, *and* $\Psi = \{\alpha \vee \beta, \neg\delta \vee \gamma\}$. *So* $\Phi_1 \preceq_\Psi \Phi_2$.

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.

Another type of analysis assumes 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.

References

Belnap, N. 1977. A useful four-valued logic. In Epstein, G., ed., *Modern Uses of Multiple-valued Logic*, 8–37. Reidel.

Benferhat, S.; Dubois, D.; Kaci, S.; and Prade, H. 2000. Encoding information fusion in possibilistic logic: A general framework for rational syntactic merging. In *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI'2000)*, 3–7. IOS Press.

Benferhat, S.; Dubois, D.; and Prade, H. 1993. Argumentative inference in uncertain and inconsistent knowledge bases. In *Proceedings of Uncertainty in Artificial Intelligence*, 1449–1445. Morgan Kaufmann.

Besnard, Ph., and Hunter, A. 1995. Quasi-classical logic: Non-trivializable classical reasoning from inconsistent information. In *Symbolic and Quantitative Approaches to Uncertainty*, volume 946 of *LNCS*, 44–51.

Brewka, G. 1989. Preferred subtheories: An extended logical framework for default reasoning. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, 1043–1048.

Cadoli, M. 1992. The complexity of model checking for circumscriptive formulae. *Information Processing Letters* 42:113–118.

da Costa, N. C. 1974. On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic* 15:497–510.

Dalal, M. 1988. Investigations into a theory of knowledge base revision: Preliminary report. In *Proceedings of the 7th National Conference on Artificial Intelligence (AAAI'88)*, 3–7. MIT Press.

Dubois, D.; Lang, J.; and Prade, H. 1994. Possibilistic logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3. Oxford University Press. 439–513.

Gardenfors, P. 1988. *Knowledge in Flux*. MIT Press.

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

Hunter, A. 2000. Reasoning with conflicting information using quasi-classical logic. *Journal of Logic and Computation* 10:677–703.

Hunter, A. 2001. A semantic tableau version of first-order quasi-classical logic. In *Symbolic and Quantitative Approaches to Uncertainty*, volume 2143 of *LNCS*, 544–556.

Kleer, J. D., and Williams, B. 1987. Diagnosing mulitple faults. *Artificial Intelligence* 32:97–130.

Konieczny, S., and Pino Perez, R. 1998. On the logic of merging. In *Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning (KR98)*, 488–498. Morgan Kaufmann.

Levesque, H. 1984. A logic of implicit and explicit belief. In *Proceedings of the National Conference on Artificial Intelligence (AAAI'84)*, 198–202.

Lin, J., and Mendelzon, A. 1998. Merging databases under constraints. *International Journal of Cooperative Information Systems* 7(1):55–76.

Lozinskii, E. 1994. Information and evidence in logic systems. *Journal of Experimental and Theoretical Artificial Intelligence* 6:163–193.

Marquis, P., and Porquet, N. 2001. Computational aspects of quasi-classical entailment. *Journal of Applied Non-classical Logics* 11:295–312.

Reiter, R. 1987. A theory of diagnosis from first principles. *Artificial Intelligence* 32:57–95.

Revesz, P. 1997. On the semantics of arbitration. *International Journal of Algebra and Computation* 7:133–160.

Schaerf, M., and Cadoli, M. 1995. Tractable reasoning via approximation. *Artificial Intelligence* 74:249–310.

Wong, P., and Besnard, Ph. 2001. Paraconsistent reasoning as an analytic tool. *Journal of the Interest Group in Propositional Logic* 9:233–246.