Reducing Preferential Paraconsistent Reasoning to Classical Entailment

Ofer Arieli
Department of Computer Science,
The Academic College of Tel-Aviv
Antokolski 4, Tel-Aviv 61161, Israel
oarieli@mta.ac.il

Marc Denecker
Department of Computer Science,
The Catholic University of Leuven
Celestijnenlaan 200A, B-3001 Heverlee, Belgium
marcd@cs.kuleuven.ac.be

August 15, 2002

Abstract

We introduce a general method for paraconsistent reasoning in the context of classical logic. A standard technique for paraconsistent reasoning on inconsistent classical theories is by shifting to multiple-valued logics. We show how these multiple-valued theories can be "shifted back" to two-valued classical theories through a polynomial transformation, and how preferential reasoning based on multiple-valued logic can be represented by classical circumscription-like axioms. By applying this process we provide new ways of implementing multiple-valued paraconsistent reasoning. Standard multiple-valued reasoning can thus be performed through theorem provers for classical logic, and multiple-valued preferential reasoning can be implemented using algorithms for processing circumscriptive theories (such as DLS and SCAN).

Keywords: paraconsistent reasoning, preferential semantics, circumscription, multiple-valued logics.

1 Introduction

It is well-known that classical logic is inappropriate for imitating "common-sense" reasoning in general, and for reasoning with uncertainty in particular. Indeed, on one hand classical logic is too cautious in drawing conclusions from incomplete theories. This is so since classical logic is monotonic, thus it does not allow to retract previously drawn conclusions in light of new, more accurate information. On the other hand, classical logic is too liberal in drawing conclusions from inconsistent theories. This is explained by the fact that classical logic is not paraconsistent [12], therefore everything classically follows from a contradictory set of premises.

Preferential reasoning [34] is an elegant way to overcome classical logic's shortcoming for reasoning on uncertainty. It is based on the idea that in order to draw conclusions from a given theory one should not consider all the models of that theory, but only a subset of *preferred models*. This subset is usually determined according to some preference criterion, which is often defined in terms of partial orders on the space of valuations. This method of preferring some models and disregarding the others yields robust formalisms that allow to draw intuitive conclusions from partial knowledge.

In the context of classical logic, preferential semantics cannot help to overcome the problem of trivial reasoning with contradictory theories. Indeed, if a certain theory has no (two-valued) models, then it has no preferred models as well. A useful way of reasoning on contradictory classical theories is therefore by embedding them in multiple-valued logics in general, and Belnap's four-valued logic [8, 9] in particular (which is the underlying multiple-valued semantics used here). There are several reasons for using this setting. The most important ones for our purposes are the following:

- In the context of four-valued semantics it is possible to define consequence relations that are not degenerated w.r.t. *any* theory (see, e.g., [3, 4, 31, 32, 35]); the fact that every theory has a nonempty set of four-valued models implies that four-valued reasoning may be useful for properly handling inconsistent theories. As shown e.g. in [3, 4], this indeed is the case.
- Analysis of four-valued models can be instructive to pinpoint the causes of the inconsistency and/or the incompleteness of the theory under consideration. (See [3, 4, 8, 9] for a detailed discussion on this property, as well as some relevant results).

However, Belnap's four-valued logic has its own shortcomings:

- As in classical logic, many theories have too many models, and as a consequence the entailment relation is often too weak. In fact, since Belnap's logic is weaker than classical logic w.r.t. consistent theories[1], we are even in a worse situation than in classical logic!

  A (partial) solution to this problem is by using preferential reasoning in the context of multiple-valued logic (see, e.g., [2, 3, 4, 5, 22, 23, 31, 32]).
- At the computational level, implementing paraconsistent reasoning based on four-valued semantics poses important challenges. An effective implementation of theorem provers for one of the existing proof systems for Belnap's logic requires a major effort. The problem is even worse in the context of four-valued *preferential* reasoning, for which currently no proof systems are known.

Our goal in this paper is to show a way in which these problems can be solved. In particular, we present a polynomial transformation back from four-valued theories to two-valued theories such that reasoning in preferential four-valued semantics can be implemented by standard theorem proving in two-valued logic.[2] Moreover, preference criteria on four-valued theories are translated into "circumscriptive-like" formulae [28, 29], and thus paraconsistent reasoning may be automatically computed by some specialized methods for compiling circumscriptive theories (such as those described in [19, 33]), and incorporated into algorithms such as SCAN [30] and DLS [13, 14], for reducing second-order formulae to their first-order equivalents.

[1] That is, everything that follows in Belnap's four-valued logic from a given theory also classically follows from that theory, but not vice-versa. For instance, the rule of excluded middle (either ψ or $\neg\psi$ should hold for every ψ) and the Disjunctive Syllogism (from $\psi \vee \phi$ and $\neg\phi$ conclude ψ) are not sound in Belnap's four-valued logic.

[2] Similar technique, of shifting back and forth between classical and non-classical logics, is also considered in [17], where it is shown that it is possible to accomplish belief revision in any logic that is translatable to classical logic.

In the last part of this paper we show that our approach of representing preferential considerations by higher-order formulae can be generalized to cases in which *arbitrarily many* truth values are needed (such as in probabilistic reasoning or fuzzy logics). For that we use Ginsberg/Fitting's *bilattices* [16, 18], which are algebraic structures that naturally extend Belnap's four-valued structure. It is shown that within the bilattice-based semantics one can use the same methods for syntactically representing preferences in many-valued logics.

The rest of this paper is organized as follows: in the next section we show how paraconsistent four-valued reasoning on a logic theory can be simulated by classical reasoning on a suitably translated first order theory. In Section 3 we show, moreover, that four-valued *preferential* reasoning on a logic theory can be simulated by amalgamating its translation with second-order formulae. In Section 4 we extend these results to general multiple-valued formalisms, and in Section 5 we conclude. [3]

2 Paraconsistent classical reasoning

In order to define the reduction of paraconsistent (four-valued) reasoning to classical logic, we first define the underlying framework for representing inconsistent (and incomplete) theories. Then we consider a polynomial transformation of theories in this framework to "*equivalent*" classical theories. Finally, in the last part of this section we use this transformation for simulating paraconsistent reasoning (as well as reasoning with incomplete information) by classical logic.

2.1 The underlying semantical structure

The formalism that we consider here is based on four-valued semantics. Four-valued reasoning may be traced back to the 1950's, where four-valued formalisms have been investigated by a number of people, including Bialynicki-Birula [10], Rasiowa [11], and Kalman [20]. In the sequel we shall use a corresponding four-valued algebraic structure (denoted here by $\mathcal{FOUR}$), introduced later by Belnap [8, 9]. This structure is composed of four elements $FOUR = \{t, f, \bot, \top\}$, arranged in the following two lattice structures:

- $(FOUR, \leq_t)$, in which t is the maximal element, f is the minimal one, and $\top, \bot$ are two intermediate and incomparable elements.
- $(FOUR, \leq_k)$, in which $\top$ is the maximal element, $\bot$ is the minimal one, and t, f are two intermediate and incomparable elements.

Here, t and f correspond to the classical truth values. The two other truth values may intuitively be understood as representing different cases of uncertainty: $\top$ corresponds to the *contradictory truth value* (i.e., the corresponding assertion and its negation are both true), and $\bot$ corresponds to an *incomplete truth value* (i.e., neither the assertion nor its negation are true). This interpretation of the meaning of the truth values will be useful in what follows for modeling paraconsistent reasoning. [4]

[3] This is a revised and an extended version of [6].

[4] This was also the original motivation of Belnap when he introduced $\mathcal{FOUR}$. See [3, 4] for some further arguments in favor of using this structure as a semantical background of formalisms for common-sense reasoning.

According to this interpretation, the partial order $\leq_t$ reflects differences in the amount of *truth* that each element represents, and the partial order $\leq_k$ reflects differences in the amount of *knowledge* (or *information*) that each element exhibits. A double-Hasse diagram of $\mathcal{FOUR}$ is given in Figure 1. [5]

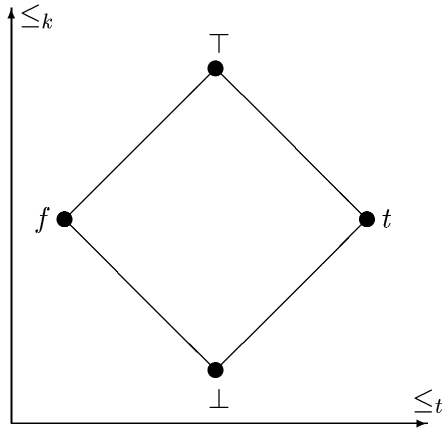

Figure 1: Belnap's four-valued structure, $\mathcal{FOUR}$

In what follows we shall denote by $\wedge$ and $\vee$ the meet and join operations on $(FOUR, \leq_t)$, and by $\otimes$ and $\oplus$ the meet and the join operations on $(FOUR, \leq_k)$. A negation, $\neg$, is a unary operation on $FOUR$, defined by $\neg t = f$, $\neg f = t$, $\neg\top = \top$, and $\neg\bot = \bot$. As usual in such cases, we take t and $\top$ as the *designated elements* in $FOUR$ (i.e., the elements that represent true assertions).

In the rest of this paper we denote by Σ a language with a finite alphabet, in which the connectives are $\vee, \wedge, \neg$. These connectives correspond to the operations on $FOUR$ with the same notations. ν and μ denote arbitrary four-valued *valuations*, i.e., functions that assign a value in $FOUR$ to every atom in Σ. The extension to complex formulae in Σ is defined in the usual way. The space of the four-valued valuations is denoted by $\mathcal{V}^4$. A valuation $\nu \in \mathcal{V}^4$ is a *model* of a formula ψ (alternatively, ν *satisfies* ψ) if $\nu(\psi) \in \{t, \top\}$. ν is a model of a set Γ of formulae if ν is a model of every $\psi \in \Gamma$. The set of the models of Γ is denoted by $mod(\Gamma)$.

2.2 An alternative representation of semantical concepts

The elements of $\mathcal{FOUR}$ can be represented by pairs of components from the two-valued lattice $(\{0,1\}, 0 < 1)$ as follows: $t = (1,0)$, $f = (0,1)$, $\top = (1,1)$, $\bot = (0,0)$. One way to understand this representation is that a four-valued truth value (x,y) for p corresponds to a two-valued truth value x for p and a two-valued truth value y for $\neg p$. Note that this reading is in accordance with the original intuitive meaning of the truth values in $\mathcal{FOUR}$, discussed in Section 2.1. According to this representation, the negation operator is defined in $\mathcal{FOUR}$ by $\neg(x,y) = (y,x)$, and the corresponding

[5]The lattices $(FOUR, \leq_t)$ and $(FOUR, \leq_k)$ are referred to in Belnap's papers [8, 9] as the *logical* lattice (L4) and the *approximation* lattice (A4), respectively, whereas the truth-values $\bot$ and $\top$ as **N**one and **B**oth, respectively. Here we follow the alternative way of denoting these elements, used, e.g., in [2, 3, 4, 5, 6, 15, 16].

partial orders in are represented by the following rules: for every $x_1, x_2, y_1, y_2 \in \{0,1\}$,

$$(x_1,y_1) \leq_t (x_2,y_2) \text{ iff } x_1 \leq x_2 \text{ and } y_1 \geq y_2, \qquad (x_1,y_1) \leq_k (x_2,y_2) \text{ iff } x_1 \leq x_2 \text{ and } y_1 \leq y_2.$$

It follows, in particular, that in the representation by pairs of two-valued components, the basic binary operations on $\mathcal{FOUR}$ are defined as follows:

$$(x_1,y_1) \vee (x_2,y_2) = (x_1 \vee x_2,\ y_1 \wedge y_2), \qquad (x_1,y_1) \wedge (x_2,y_2) = (x_1 \wedge x_2,\ y_1 \vee y_2),$$

$$(x_1,y_1) \oplus (x_2,y_2) = (x_1 \vee x_2,\ y_1 \vee y_2), \qquad (x_1,y_1) \otimes (x_2,y_2) = (x_1 \wedge x_2,\ y_1 \wedge y_2).$$

It is obvious that the above representation of the truth values in terms of pairs of two-valued components implies a similar way of representing four-valued valuations; a four-valued valuation ν may represented in terms of pair of two-valued components (ν_1, ν_2) by $\nu(p) = (\nu_1(p), \nu_2(p))$. So if, for instance, $\nu(p) = t$, then $\nu_1(p) = 1$ and $\nu_2(p) = 0$.

Next we propose a technique to compute the truth value of a formula in four-valued logic, by transforming it to a formula that can be evaluated in two-valued logic. This is called a *splitting transformation*.

Define the scope of a negation operator $\neg$ in the formula $\neg\psi$ as the set of all appearances of propositional symbols in ψ.

Definition 2.1 Let ψ be a formula. We say that an appearance of p in ψ is *positive*, if it appears in the scope of an even number of negation operators in ψ; otherwise, it is a *negative* appearance.

Example 2.2 Let $\psi = \neg(p \vee \neg q) \vee \neg q$. Then the first appearance of q in ψ is positive, and the second appearance of q in ψ is negative.

Let ψ be a formula in a language Σ. Denote by $\overline{\psi}$ the formula that is obtained by substituting every positive occurrence in ψ of an atomic formula p by a new predicate symbol p^+, and replacing every negative occurrence in ψ of an atomic formula p by $\neg p^-$. For instance, if ψ is the formula of Example 2.2, then $\overline{\psi} = \neg(\neg p^- \vee \neg q^+) \vee \neg\neg q^-$. Given a theory Γ, we shall write $\overline{\Gamma}$ for the set $\{\overline{\psi} \mid \psi \in \Gamma\}$. The language that is obtained from Σ by introducing these new predicate symbols will be denoted in what follows by $\Sigma^\pm$.

Given a four-valued valuation $\nu = (\nu_1, \nu_2)$ of the atomic formulae $\vec{p}$ in Σ, $\overline{\nu}$ denotes the two-valued valuation on the atoms of $\Sigma^\pm$, which interprets symbols p^+ as $\nu_1(p)$ and symbols p^- as $\nu_2(p)$. Sometimes we shall "unfold" this notation and instead of $\overline{\nu}$ we shall write $(\vec{p}^{\,+}\!:\!\nu_1\ ;\ \vec{p}^{\,-}\!:\!\nu_2)$. $\overline{\nu}$ is a standard two-valued interpretation of atoms in $\Sigma^\pm$ and can be extended to complex formulae in the usual way. [6]

Notation 2.3 For a valuation $\overline{\nu} = (\vec{p}^{\,+}\!:\!\nu_1\ ;\ \vec{p}^{\,-}\!:\!\nu_2)$, denote $\texttt{rev}(\overline{\nu}) = (\vec{p}^{\,+}\!:\!\neg\nu_2\ ;\ \vec{p}^{\,-}\!:\!\neg\nu_1)$.

[6] Clearly, the converse construction is also possible: every two-valued valuation ν on $\Sigma^\pm$ corresponds to a unique four-valued valuation ν' on Σ defined, for every atom p, by $\nu'(p) = (\nu(p^+), \nu(p^-))$.

Proposition 2.4 Let $\nu = (\nu_1, \nu_2)$. Then $\nu(\psi) = (\overline{\nu}(\overline{\psi}),\ \neg\texttt{rev}(\overline{\nu})(\overline{\psi}))$.

Proof: The proof is rather technical; we give it in the appendix. $\square$

Note 2.5 Proposition 2.4 shows how to represent valuations in terms of their split counterparts. It will have a central role in showing some of the next results (e.g., Theorems 2.8 and 4.15 below). It is interesting to note that this proposition holds also w.r.t. another (polynomial) splitting transformation of formulae in the language, used e.g. in [6], which can be defined as follows: for a formula ψ in Σ denote by $\widetilde{\psi}$ the formula in $\Sigma^{\pm}$, obtained from ψ by first translating ψ to its negation normal form, ψ' (where the negation operator precedes atomic formulae only),[7] then substituting every atomic formula p, which is not preceded by a negation, by a new predicate symbol p^+, and replacing every other atomic formula q in ψ', together with the negation that precedes it, by a new predicate symbol q^-. For instance, if $\psi = \neg(r \vee \neg s)$, then $\widetilde{\psi} = r^- \wedge s^+$.

Proposition 2.4 holds also w.r.t. the alternative transformation (see [6] for the proof), and so all the relevant results in the sequel can be obtained w.r.t. either one of the splitting transformations. Here we shall use the original transformation, which is applied directly on the underlying formula rather than on its representation in negation normal form.

2.3 Simulating four-valued reasoning by classical logic

A natural definition of a consequence relation on $\mathcal{FOUR}$ is the following:

Definition 2.6 Let Γ be a set of formulae and ψ a formula in Σ. Denote $\Gamma \models^4 \psi$ if every four-valued model of Γ is a four-valued model of ψ.

In [2] it is shown that $\models^4$ is a consequence relation in the sense of Tarski [36], and that it is paraconsistent and compact. Moreover, as it is shown in [2, 7], $\models^4$ has a cut-free, sound and complete Gentzen-type proof system (known as "the $\{\wedge, \vee, \neg\}$-fragment of *GBL*" in [2], or "the basic $\{\wedge, \vee, \neg\}$-system" in [7]), which is the same as the system of "first degree entailments" in relevance logic [1]. In fact, for formulae $\psi, \psi_1, \ldots, \psi_n$ in Σ, the following conditions are equivalent:

1. $\psi_1, \ldots, \psi_n \models^4 \psi$.
2. $\psi_1, \ldots, \psi_n \Rightarrow \psi$ is provable in *GBL* (or in the "intuitionistic" version of it, GBL_I [2])
3. $\psi_1 \wedge \ldots \wedge \psi_n \rightarrow \psi$ is provable in the system R (or E) of Anderson and Belnap [1].
4. For every valuation ν in $\mathcal{V}^4$, $\nu(\psi_1 \wedge \ldots \wedge \psi_n) \leq_t \nu(\psi)$.

In particular, the last item provides an alternative definition for $\models^4$ w.r.t. the language Σ (cf. Definition 2.6).

In what follows we use the pairwise representations, considered in Section 2.2, for showing that four-valued reasoning can be simulated by classical reasoning. The justification for doing so is the following lemma:

[7] It is easy to verify that as in the two-valued case, ψ and ψ' are logically equivalent in $\mathcal{FOUR}$.

Lemma 2.7 For every four-valued valuation ν and a formula ψ in Σ, $\nu(\psi)$ is designated iff $\overline{\nu}(\overline{\psi})=1$.

Proof: $\nu(\psi)$ is designated iff $\nu_1(\psi)=1$, iff (Proposition 2.4) $\overline{\nu}(\overline{\psi})=1$. □

The following result is an immediate corollary of Lemma 2.7:

Theorem 2.8 Denote by $\models^2$ the two-valued classical consequence relation. Then $\Gamma \models^4 \psi$ iff $\overline{\Gamma} \models^2 \overline{\psi}$.

It follows, therefore, that four-valued reasoning may be implemented by two-valued theorem provers. Moreover, since $\overline{\Gamma}$ is obtained from Γ in a polynomial time, the theorem above shows that four-valued entailment in the context of Belnap's logic is *polynomially reducible* to classical entailment.

Another immediate consequence of this theorem is the next well-known result:

Corollary 2.9 In positive logic (i.e., in the language without negations), $\Gamma \models^4 \psi$ iff $\Gamma \models^2 \psi$.

Proof: Follows from Theorem 2.8 and the fact that in the language without negations $\overline{\Gamma}$ is the same as Γ. □

Note 2.10 Recently and independently, Gabbay, Rodrigues and Russo [17] have provided another way (motivated by different arguments) to simulate (basic) four-valued reasoning by classical entailment. Roughly, the idea in [17] is to use a two sorted first-order language, composed of a sort for representing formulae (in the language of $\wedge, \vee, \neg$) and a sort for the truth-values. The language also contains a special two-sorted binary predicate, denoted *hold*, where $holds(\psi, \mathtt{t})$ intuitively means that there is some $\nu \in \mathcal{V}^4$ s.t. $\nu(\psi) \geq_k t$, and $holds(\psi, \mathtt{f})$ intuitively means that there is some $\nu \in \mathcal{V}^4$ s.t. $\nu(\psi) \geq_k f$. In addition, a set $\mathcal{A}$ with the following axioms is considered:

$\forall x \; (\, holds(x, \mathtt{f}) \leftrightarrow holds(\neg x, \mathtt{t}) \,)$	$\forall x \; (\, holds(x, \mathtt{t}) \leftrightarrow holds(\neg x, \mathtt{f}) \,)$
$\forall x \forall y \; (\, holds(x \wedge y, \mathtt{t}) \leftrightarrow holds(x, \mathtt{t}) \wedge holds(y, \mathtt{t}) \,)$	$\forall x \forall y \; (\, holds(x \wedge y, \mathtt{f}) \leftrightarrow holds(x, \mathtt{f}) \vee holds(y, \mathtt{f}) \,)$
$\forall x \forall y \; (\, holds(x \vee y, \mathtt{t}) \leftrightarrow holds(x, \mathtt{t}) \vee holds(y, \mathtt{t}) \,)$	$\forall x \forall y \; (\, holds(x \vee y, \mathtt{f}) \leftrightarrow holds(x, \mathtt{f}) \wedge holds(y, \mathtt{f}) \,)$

In the present notations, the following result holds (cf. Theorem 2.8):

Proposition 2.11 [17] Let Γ be a set of formulae and ψ a formula in Σ. Then:
$\Gamma \models^4 \psi$ iff $\mathcal{A} \cup \{holds(\gamma, \mathtt{t}) \mid \gamma \in \Gamma\} \models^2 holds(\psi, \mathtt{t})$ and $\mathcal{A} \cup \{\neg holds(\gamma, \mathtt{f}) \mid \gamma \in \Gamma\} \models^2 \neg holds(\psi, \mathtt{f})$.

3 Preferential reasoning

Despite the nice properties of $\models^4$, it appears that it has several drawbacks. One of which is that $\models^4$ is strictly weaker than classical logic, even for consistent theories (e.g., $\not\models^4 \psi \vee \neg\psi$). Also, it completely invalidates some intuitively justified inference rules, like the Disjunctive Syllogism: from $\neg\psi$ and $\psi \vee \phi$ one cannot infer ψ by using $\models^4$. Finally, the fact that $\models^4$ is a Tarskian consequence relation means, in particular, that it is monotonic, and as such it is "over-cautious" in drawing conclusions from incomplete theories.

In order to overcome these drawbacks of $\models^4$, we consider in this section a refined way of drawing consequences from a given theory, known as *preferential reqsoning* [34]. This is a general model theory for non-monotonic inferences, in which the set of the semantical objects that describe a given theory is equipped with a preference relation that intuitively reflects some preference criterion among the given semantical objects. Inferences are then made only according to those elements that are the most-preferred ones w.r.t. the preference relation.

In the first part of this section we quickly revise some basic concepts and notations that are related to preferential reasoning.[8] Then we recall the techniques of [4] for applying preferential reasoning in the four-valued context. In the third part of this section we show that, again, classical logic can be used for simulating the kind of reasoning in a four-valued semantics that we are interested in. This time, in addition to the basic theory that is converted to a classical one, (second-order) circumscribing formulae will be used for representing the corresponding preference relations. We conclude this section with experimental results for some simple test cases.

3.1 Preliminaries

Definition 3.1 A *preferential model* (w.r.t. a language Σ) is a triple $\mathcal{M} = (M, \models, \leq)$, where

- M is a set (of semantical objects, sometimes called *states*),
- $\models$ is a relation on $M \times \Sigma$ (sometimes called the *satisfaction relation*), and
- $\leq$ is a binary relation on the elements of M (sometimes called the *preference relation*).

Note that Definition 3.1 is a very general one. Some formalisms make more specific assumptions on the nature of the components of a preferential model. For instance, in the original definition of Shoham [34], each preferential model corresponds to a theory Γ, the underlying semantical objects (i.e., the elements in M) are the models of Γ w.r.t. the satisfaction relation $\models$, and the preferential relation $\leq$ is a partial order on M.

Definition 3.2 Let $\mathcal{M} = (M, \models, \leq)$ be a preferential model, Γ a set of formulae in a language Σ, and $m \in M$.

a) m *satisfies* Γ (notation: $m \models \Gamma$) if $m \models \gamma$ for every $\gamma \in \Gamma$.

b) m *preferentially satisfies* Γ (alternatively, m is a $\leq$*-most preferred model* of Γ) if m satisfies Γ and there is *no element* $n \in M$ that satisfies Γ, and for which $n \leq m$ and $m \not\leq n$.

c) The set of the elements in M that preferentially satisfy Γ is denoted by $!(\Gamma, \leq)$.

Now we can define the preferential entailment relations:

Definition 3.3 Let $\mathcal{M} = (M, \models, \leq)$ be a preferential model, Γ a set of formulae in Σ, and ψ a formula in Σ. We say that ψ (preferentially) *follows* from Γ (alternatively, Γ preferentially *entails* ψ), if every element of $!(\Gamma, \leq)$ satisfies ψ. We denote this by $\Gamma \models_{\leq} \psi$.

[8] Among the various ways of defining preferential reasoning that are given in the literature, we follow here that of Makinson [27].

In case that M consists of the models of Γ, Definition 3.3 simply says that Γ preferentially entails ψ if every $\leq$-preferred model of Γ is a model of ψ.

The idea that a non-monotonic deduction should be based on some preference criterion that reflects some normality relation among the relevant semantical objects is a very natural one, and may be traced back to [28]. Furthermore, this approach is the semantical basis of some well-known general patterns for non-monotonic reasoning, introduced in [24, 25, 26, 27], and it is a key concept behind many formalisms for nonmonotonic and paraconsistent reasoning, such as RI [22, 23], LPm [31, 32], and the bilattice-based logics of [2, 5]. Our purpose in the rest of this paper is to propose techniques of expressing some of the preferential relations used in these formalisms by formulae in the underlying language. Next we define the framework for doing so.

3.2 Four-valued preferential reasoning

In what follows we describe some particularly useful ways of applying preferential reasoning in the four-valued case. See [4] for a more detailed discussion on the formailsms that are obtained.

Definition 3.4 [4] Let $\nu, \mu \in \mathcal{V}^4$. Denote:

- $\nu \leq_k \mu$ if $\nu(p) \leq_k \mu(p)$ for every atom p.
- $\nu \leq_{\{\top\}} \mu$ if for every atom p, $\mu(p) = \top$ whenever $\nu(p) = \top$.
- $\nu \leq_{\{\top,\bot\}} \mu$ if for every atom p, $\mu(p) \in \{\top, \bot\}$ whenever $\nu(p) \in \{\top, \bot\}$.

It is easy to check that $\leq_k$ is a partial order and $\leq_{\{\top\}}$, $\leq_{\{\top,\bot\}}$ are pre-orders on $\mathcal{V}^4$. In what follows we shall write $\nu <_k \mu$ to denote that $\nu \leq_k \mu$ and $\mu \not\leq_k \nu$; similarly for $<_{\{\top\}}$ and $<_{\{\top,\bot\}}$.

Each one of these preference orders has its own rationality: according to $\leq_k$, for instance, one prefers valuations with as minimal information as reasonably possible. This is a common criterion for making preferences among different semantics of a given theory.[9] This criterion may as well be viewed as an argumentation for consistency preserving, since as long as one keeps the amount of information (or belief) as minimal as possible, the tendency of getting into conflicts decreases.

The pre-order $\leq_{\{\top\}}$ states a somewhat more explicit preference of inconsistency minimization: it prefers those valuations that minimize the amount of inconsistent assignments. The last order given in Definition 3.4, $\leq_{\{\top,\bot\}}$, prefers those valuations that are as classical as possible, i.e., those ones that assign classical truth values whenever possible.

In terms of Section 2.2, the preference criteria of Definition 3.4 may be reformulated as follows:

Lemma 3.5 Let $\nu, \mu \in \mathcal{V}^4$. Then:

- $\nu \leq_k \mu$ iff for every atom p, $\nu_1(p) \leq \mu_1(p)$ and $\nu_2(p) \leq \mu_2(p)$.

[9] Notable examples of formalisms that are based on the idea of $\leq_k$-minimization are the well-founded semantics [37] and Fitting's fixpoint semantics [15] for general logic programs.

- $\nu \leq_{\{\top\}} \mu$ iff for every atom p, whenever $\nu_1(p) \wedge \nu_2(p) = 1$, $\mu_1(p) \wedge \mu_2(p) = 1$ as well.
- $\nu \leq_{\{\top,\perp\}} \mu$ iff for every atom p, whenever $(\nu_1(p) \wedge \nu_2(p)) \vee (\neg\nu_1(p) \wedge \neg\nu_2(p)) = 1$, we have that $(\mu_1(p) \wedge \mu_2(p)) \vee (\neg\mu_1(p) \wedge \neg\mu_2(p)) = 1$ as well.

Proof: Immediately follows from the corresponding definitions. □

Alternatively, the preference criteria above may be defined as follows:

Corollary 3.6 Let $\nu, \mu \in \mathcal{V}^4$. Then:

- $\nu \leq_k \mu$ iff whenever $\nu_i(p) = 1$ then $\mu_i(p) = 1$, for $i = 1, 2$.
- $\nu \leq_{\{\top\}} \mu$ iff whenever $\nu_1(p) = \nu_2(p) = 1$ then $\mu_1(p) = \mu_2(p) = 1$.
- $\nu \leq_{\{\top,\perp\}} \mu$ iff whenever $\nu_1(p) = \nu_2(p)$ then $\mu_1(p) = \mu_2(p)$.

Given a set Γ of formulae in Σ, the minimal elements in $mod(\Gamma)$ w.r.t. $\leq_k$ are called the *k-minimal models* of Γ.[10] Similarly, the minimal elements of $mod(\Gamma)$ w.r.t. $\leq_{\{\top\}}$ are called the *most consistent models* of Γ, and the minimal elements of $mod(\Gamma)$ w.r.t. $\leq_{\{\top,\perp\}}$ are called the *most classical models* of Γ.

Example 3.7 Let $\Gamma = \{p, \neg p \vee q, \neg q, r \vee q\}$. The ten four-valued models of Γ are given in Table 1.

Table 1: The elements in $mod(\Gamma)$

Model No.	p	q	r
$M_1 - M_2$	$\top$	f	$t, \top$
$M_3 - M_6$	t	$\top$	$\perp, f, t, \top$
$M_7 - M_{10}$	$\top$	$\top$	$\perp, f, t, \top$

Thus, the k-minimal models of Γ are $\{M_1, M_3\}$, the most consistent ones are $\{M_1, M_3, M_4, M_5\}$, and the most classical ones are $\{M_1, M_4, M_5\}$.

Each one of the preference criteria considered in Definition 3.4 induces a corresponding preferential consequence relation. Next we define these relations:

Definition 3.8 [4] Let Γ be a set of formulae and ψ a formula in Σ. Denote:

- $\Gamma \models^4_k \psi$ if every k-minimal model of Γ is a model of ψ.
- $\Gamma \models^4_{\{\top\}} \psi$ if every most consistent model of Γ is a model of ψ.
- $\Gamma \models^4_{\{\top,\perp\}} \psi$ if every most classical model of Γ is a model of ψ.

[10] That is, $\nu \in mod(\Gamma)$ is a k-minimal model of Γ if there is no $\mu \in mod(\Gamma)$ s.t. $\mu <_k \nu$.

Example 3.9 Consider again the set Γ of Example 3.7, and let $\psi = r \vee \neg r$. Then $\Gamma \models^4_{\{\top,\bot\}} \psi$, while $\Gamma \not\models^4_k \psi$ and $\Gamma \not\models^4_{\{\top\}} \psi$.

Clearly, the consequence relations defined in 3.8 are particular cases of the preferential entailment relations $\models_{\leq}$, defined in 3.3 (see also the note after Definition 3.3). Some important properties of these relations are listed in the next proposition:

Proposition 3.10 [4] Denote by $\models^2$ the two-valued classical consequence relation. For every set Γ of formulae and a formula ψ in Σ,

1. $\Gamma \models^4_k \psi$ iff $\Gamma \models^4 \psi$.
2. If Γ is classically consistent and ψ is a formula in CNF, none of its disjunctions is a tautology, then $\Gamma \models^4_{\{\top\}} \psi$ iff $\Gamma \models^2 \psi$.
3. If Γ is classically consistent then $\Gamma \models^4_{\{\top,\bot\}} \psi$ iff $\Gamma \models^2 \psi$.
4. $\models^4_k$, $\models^4_{\{\top\}}$, and $\models^4_{\{\top,\bot\}}$ are paraconsistent.

Note 3.11 Proposition 3.10 demonstrates the usefulness of the consequence relations considered in Definition 3.8:

- Item 1 implies that $\models^4_k$ is a compact representation of $\models^4$; it is sufficient to consider only the k-minimal models of a given theory in order to simulate reasoning with $\models^4$.
- By item 2 it follows that in order to check whether a formula classically follows from a consistent theory Γ, it is sufficient to convert it to a conjunctive normal form, drop all the conjuncts that are tautologies, and check the remaining formula only w.r.t. the most consistent models of Γ.
- By items 3 and 4 it follows that $\models^4_{\{\top,\bot\}}$ is equivalent to classical logic on consistent theories and is nontrivial w.r.t. inconsistent theories.

A more detailed discussion on the consequence relations defined in 3.8 and some related ones appears in [4, 5]. In the next section we will show how to express the semantical considerations behind such relations by second-order formulae.

3.3 Simulating preferential four-valued reasoning by circumscription

In this section we show that four-valued preferential entailment can be defined in terms of classical entailment for the transformed theories augmented with circumscriptive axioms. Indeed, in order to extend the technique of Section 2.3 to deal with preferential four-valued reasoning, we must express that the encoded four-valued interpretation is minimal with respect to the preference relation $\leq$. This is accomplished by introducing a circumscription axiom that expresses the preference relation $\leq$ objectively, by a formula $\Psi_{\leq}$. Thus, the first point to check out is how to express a semantical preference relation $\leq$ in an axiom.

Let $\vec{p} = \{p_1, p_2, \ldots\}$ be the set symbols of a language Σ. Define $\vec{p}^{\,\pm}$ as the set of symbols $\{p_1^+, p_1^-, p_2^+, p_2^-, \ldots\}$. To be able to express for two valuations $\nu = (\nu_1, \nu_2)$ and $\mu = (\mu_1, \mu_2)$ that $\nu \leq \mu$ by one formula, we introduce new symbols $\vec{q}$ as a renaming of the symbols of $\vec{p}$. Similar as before, we define $(\vec{p}^{\,\pm} : \nu\, ;\, \vec{q}^{\,\pm} : \mu)$ as the two-valued interpretation that interprets, for every i, the symbols p_i^+ as $\nu_1(p)$, p_i^- as $\nu_2(p)$, q_i^+ as $\mu_1(p)$ and q_i^- as $\mu_2(p)$.

Definition 3.12 A preferential order $\leq$ is *represented* by a formula $\Psi_{\leq}(\vec{p}^{\,\pm}, \vec{q}^{\,\pm})$ if for every four-valued valuations ν and μ we have that $\nu \leq \mu$ iff $(\vec{p}^{\,\pm} : \nu,\, \vec{q}^{\,\pm} : \mu)$ satisfies $\Psi_{\leq}(\vec{p}^{\,\pm}, \vec{q}^{\,\pm})$.

Proposition 3.13 Let $\Psi_{\leq}(\vec{p}^{\,\pm}, \vec{q}^{\,\pm})$ be a formula that represents a preferential order $\leq$. Then ν is a $\leq$-most preferred model of ψ (that is, $\nu \in\, !(\psi, \leq)$) iff $\overline{\nu}$ satisfies $\overline{\psi}$ and the following formula:

$$\mathsf{Circ}_{\leq}(\vec{p}^{\,\pm}) \;=\; \forall(\vec{q}^{\,\pm})\, \{\, \overline{\psi}(\vec{q}^{\,\pm}) \;\rightarrow\; (\, \Psi_{\leq}(\vec{q}^{\,\pm}, \vec{p}^{\,\pm}) \rightarrow \Psi_{\leq}(\vec{p}^{\,\pm}, \vec{q}^{\,\pm})\,)\, \}.$$

Proof: By Corollary 2.7, ν is a model of ψ iff $\overline{\nu}$ satisfies $\overline{\psi}$. It remains to show that the fact that $\overline{\nu}$ satisfies $\mathsf{Circ}_{\leq}$ is a necessary and sufficient condition for assuring that ν is a $\leq$-minimal element in the set $mod(\psi)$ of the models of ψ. Indeed, $\overline{\nu}$ satisfies $\mathsf{Circ}_{\leq}$ iff for every valuation μ that satisfies ψ and for which $\mu \leq \nu$, it is also true that $\nu \leq \mu$. Thus, for every $\mu \in mod(\psi)$, we have that $(\mu \leq \nu) \rightarrow (\nu \leq \mu)$ (alternatively, there is no $\mu \in mod(\psi)$ s.t. $\mu < \nu$), i.e., $\nu \in\, !(\psi, \leq)$. $\square$

Note 3.14 Let $\Psi_{<}(\vec{q}^{\,\pm}, \vec{p}^{\,\pm}) = \Psi_{\leq}(\vec{q}^{\,\pm}, \vec{p}^{\,\pm}) \wedge \neg \Psi_{\leq}(\vec{p}^{\,\pm}, \vec{q}^{\,\pm})$,[11] and denote by $\vec{p}^{\,\pm} = \vec{q}^{\,\pm}$ the formula $\bigwedge_{i=1}^{n} ((p_i^+ = q_i^+) \wedge (p_i^- = q_i^-))$. Then

a) The formula $\mathsf{Circ}_{\leq}$ of Proposition 3.13 may be rewritten as follows:

$$\mathsf{Circ}_{\leq}(\vec{p}^{\,\pm}) \;=\; \forall(\vec{q}^{\,\pm})\; \{\; \overline{\psi}(\vec{q}^{\,\pm}) \;\rightarrow\; \neg \Psi_{<}(\vec{q}^{\,\pm}, \vec{p}^{\,\pm})\; \}$$

b) In case that $\leq$ is a partial order, $\mathsf{Circ}_{\leq}$ can be rewritten as follows:

$$\mathsf{Circ}_{\leq}(\vec{p}^{\,\pm}) \;=\; \forall(\vec{q}^{\,\pm})\, \{\, [\, \overline{\psi}(\vec{q}^{\,\pm}) \wedge \Psi_{\leq}(\vec{q}^{\,\pm}, \vec{p}^{\,\pm})\,] \rightarrow \vec{p}^{\,\pm} = \vec{q}^{\,\pm}\, \}$$

The next theorem is an immediate corollary of Proposition 3.13:

Notation 3.15 Denote by $\models^4_{\leq}$ the consequence relation $\models_{\leq}$ (Definition 3.3), where the underlying semantical structure is $\mathcal{FOUR}$, and the set of the designated elements is $\mathcal{D} = \{t, \top\}$.

Theorem 3.16 Let Γ be a set of formula and ψ a formula in Σ. Let $\mathsf{Circ}_{\leq}$ be the formula given in Proposition 3.13 for a preferential relation $\leq$. Then $\Gamma \models^4_{\leq} \psi$ iff $\overline{\Gamma} \cup \mathsf{Circ}_{\leq} \models^2 \overline{\psi}$.

Proposition 3.13 gives a general characterization in terms of "formula circumscription" [29] of the preferred models of a given theory: given a preferential relation $\leq$, in order to express $\leq$-preferential satisfaction of a theory, one should first formulate a corresponding formula $\Psi_{\leq}$ that represents $\leq$, and then integrate $\Psi_{\leq}$ with $\mathsf{Circ}_{\leq}$ as in Proposition 3.13. Again, this can be done in a polynomial time.

Next we define formulae that represent the preferential relations considered above. For that, we shall need the following notations:

[11] It is easy to see that for all four-valued valuations μ and ν, $\mu < \nu$ iff $(\vec{q}^{\,\pm} : \mu,\, \vec{p}^{\,\pm} : \nu)$ satisfies $\Psi_{<}(\vec{q}^{\,\pm}, \vec{p}^{\,\pm})$.

Notation 3.17 In what follows we shall write $x \preceq y$ for $x \to y$, and $x \prec y$ for $(x \to y) \wedge \neg(y \to x)$.[12]

Lemma 3.18 Let n be the number of different atomic formulae in Σ. Then:

a) The preferential relation $\leq_k$ is represented by the following formula:

$$\Psi_{\leq_k}(\vec{p}^{\,\pm}, \vec{q}^{\,\pm}) \;=\; \bigwedge_{i=1}^{n} ((p_i^+ \preceq q_i^+) \wedge (p_i^- \preceq q_i^-)).$$

b) The preferential relation $\leq_{\{\top\}}$ is represented by the following formula:

$$\Psi_{\leq_{\{\top\}}}(\vec{p}^{\,\pm}, \vec{q}^{\,\pm}) \;=\; \bigwedge_{i=1}^{n} ((p_i^+ \wedge p_i^-) \;\preceq\; (q_i^+ \wedge q_i^-)).$$

c) The preferential relation $\leq_{\{\top,\bot\}}$ is represented by the following formula:

$$\Psi_{\leq_{\{\top\}}}(\vec{p}^{\,\pm}, \vec{q}^{\,\pm}) \;=\; \bigwedge_{i=1}^{n} (\,((p_i^+ \wedge p_i^-) \vee (\neg p_i^+ \wedge \neg p_i^-)) \;\preceq\; ((q_i^+ \wedge q_i^-) \vee (\neg q_i^+ \wedge \neg q_i^-))\,).$$

Proof: We show only part (a); the proof of the other parts is similar.

$$\begin{aligned}
\nu \leq_k \mu \;&\Longleftrightarrow\; \forall 1 \leq i \leq n \;\; \nu(p_i) \leq_k \mu(p_i) \\
&\Longleftrightarrow\; \forall 1 \leq i \leq n \;\; \nu_1(p_i) \leq \mu_1(p_i) \text{ and } \nu_2(p_i) \leq \mu_2(p_i) \\
&\Longleftrightarrow\; (\vec{p}^{\,\pm} : \nu,\; \vec{q}^{\,\pm} : \mu) \text{ satisfies } \textstyle\bigwedge_{i=1}^{n} ((p_i^+ \preceq q_i^+) \wedge (p_i^- \preceq q_i^-)) \\
&\Longleftrightarrow\; (\vec{p}^{\,\pm} : \nu,\; \vec{q}^{\,\pm} : \mu) \text{ satisfies } \Psi_{\leq_k}(\vec{p}^{\,\pm}, \vec{q}^{\,\pm}) \qquad \square
\end{aligned}$$

By Proposition 3.13, Lemma 3.18(a), and Note 3.14(b), we have the following corollary:

Corollary 3.19 A valuation $\nu = (\nu_1, \nu_2)$ is a k-minimal model of ψ iff $\overline{\nu}$ satisfies $\overline{\psi}$ and $\mathsf{Circ}_{\leq_k}(\vec{p}^{\,\pm})$, where $\mathsf{Circ}_{\leq_k}(\vec{p}^{\,\pm})$ is the following formula: [13]

$$\forall(\vec{q}^{\,\pm})\;\{\;[\;\overline{\psi}(\vec{q}^{\,\pm}) \;\wedge\; \bigwedge_{i=1}^{n} ((q_i^+ \preceq p_i^+) \wedge (q_i^- \preceq p_i^-))\;] \;\to\; [\;\bigwedge_{i=1}^{n} ((q_i^+ = p_i^+) \wedge (q_i^- = p_i^-))\;]\;\}$$

As in Corollary 3.19, the most consistent models and the most classical models of a given theory can be represented by formulae of the form $\mathsf{Circ}_{\leq_{\{\top\}}}(\vec{p}^{\,\pm})$ and $\mathsf{Circ}_{\leq_{\{\top,\bot\}}}(\vec{p}^{\,\pm})$, obtained by respectively integrating the formulae given in parts (b) and (c) of Lemma 3.18 with $\mathsf{Circ}_{\leq}$, given in Proposition 3.13.

In what follows we consider a uniform way of representing $\mathsf{Circ}_{\leq_{\{\top\}}}(\vec{p}^{\,\pm})$, $\mathsf{Circ}_{\leq_{\{\top,\bot\}}}(\vec{p}^{\,\pm})$, and some other formulae that correspond to preferential criteria like $\leq_{\{\top\}}$ and $\leq_{\{\top,\bot\}}$. For this, let

[12] Thus, $x \prec y = (x \preceq y) \wedge \neg(y \preceq x)$.

[13] Note that $\mathsf{Circ}_{\leq_k}(\vec{p}^{\,\pm})$ is a standard circumscriptive axiom in the sense of [28].

$\Delta \subseteq FOUR$. Define an order relation $<_\Delta$ on $FOUR$ by $x <_\Delta y$ iff $x \notin \Delta$ while $y \in \Delta$. A corresponding pre-order on $\mathcal{V}^4$ may now be defined as follows: for every $\nu, \mu \in \mathcal{V}^4$, $\nu \leq_\Delta \mu$ iff for every atom p, the fact that $\nu(p) \in \Delta$ entails that $\mu(p) \in \Delta$ as well. The $\leq_\Delta$-most preferred models of Γ are the $\leq_\Delta$-minimal elements in $mod(\Gamma)$, and $\Gamma \models^4_\Delta \psi$ if every $\leq_\Delta$-most preferred model of Γ is a model of ψ.

Clearly, $\leq_{\{\top\}}$ and $\leq_{\{\top,\bot\}}$ are particular cases of $\leq_\Delta$, where $\Delta = \{\top\}$ and $\Delta = \{\top, \bot\}$, respectively. Now, the $\leq_\Delta$-most preferred models of a given theory can be represented by a circumscriptive formula in the following way:

Notation 3.20 For $\Delta \subseteq FOUR$, let $\Lambda_\Delta(p^+,p^-) = \bigvee_{x\in\Delta} \Lambda_x(p^+,p^-)$, where $\Lambda_t(p^+,p^-) = p^+ \wedge \neg p^-$, $\Lambda_f(p^+,p^-) = \neg p^+ \wedge p^-$, $\Lambda_\bot(p^+,p^-) = \neg p^+ \wedge \neg p^-$, and $\Lambda_\top(p^+,p^-) = p^+ \wedge p^-$. [14]

Similar arguments as those in Lemma 3.18 show that the formula

$$\Psi_{\leq_\Delta}(\vec{p}^{\,\pm},\vec{q}^{\,\pm}) \;=\; \bigwedge_{i=1}^{n} (\, \Lambda_\Delta(p_i^+,p_i^-) \preceq \Lambda_\Delta(q_i^+,q_i^-)\,)$$

represents the preferential relation $\leq_\Delta$. Therefore, by Proposition 3.13,

Proposition 3.21 A valuation ν is a $\leq_\Delta$-preferred model of ψ iff $\overline{\nu}$ satisfies $\overline{\psi}$ and the following formula:

$$\mathsf{Circ}_{\leq_\Delta}(\vec{p}^{\,\pm}) \;=\; \forall(\vec{q}^{\,\pm})\,\{\, \overline{\psi}(\vec{q}^{\,\pm}) \;\rightarrow\; (\,\Psi_{\leq_\Delta}(\vec{q}^{\,\pm},\vec{p}^{\,\pm}) \rightarrow \Psi_{\leq_\Delta}(\vec{p}^{\,\pm},\vec{q}^{\,\pm})\,)\,\}.$$

3.4 Experimental study

As we have already noted, all the formulae that are obtained by our method have a circumscriptive form. It is therefore possible to apply, for instance, the formula $\mathsf{Circ}_{\leq_k}$, given in Corollary 3.19, in algorithms for reducing circumscriptive axioms. Below are some simple results obtained by experimenting with such algorithm (We have used Doherty, Lukaszewicz and Szalas DLS algorithm [13, 14], available at `http://www.ida.liu.se/labs/kplab/projects/dls/circ.html`). [15]

- Consider the theory $\Gamma = \{Q(a), Q(b), \neg Q(a)\}$, where Q denotes some predicate, and a, b are two constants. In our context, this theory is translated to $\overline{\Gamma} = \{Q^+(a), Q^+(b), Q^-(a)\}$. Circumscribing $\overline{\Gamma}$ where Q^+ and Q^- are simultaneously minimized, yields the following result:

  $$\forall x\, \{\, (Q^-(x) \rightarrow x = a) \;\wedge\; (Q^+(x) \rightarrow (x = a \vee x = b))\,\}.$$

  It follows, then, that a is the only object for which both $Q^+(x)$ and $Q^-(x)$ hold (i.e., a is the only object that is inconsistent w.r.t. Q), and b is the only object for which only $Q^+(x)$ holds. For all the other objects neither $Q^+(x)$ nor $Q^-(x)$ holds, i.e., if $c \notin \{a, b\}$ then $Q(c)$ corresponds to $\bot$. This indeed is exactly the k-minimal semantics of Γ.

  Note that the fact that for every object x different from a and b neither $Q^+(x)$ nor $Q^-(x)$ holds means that the truth values of all the domain elements other than a or b do not matter in order to satisfy this formula. This may be important from an analytic point of view.

[14] Intuitively, $\Lambda_x(p^+,p^-)$ expresses that $\nu(p) = x$ and $\Lambda_\Delta(p^+,p^-)$ expresses that $\nu(x) \in \Delta$.

[15] In what follows we deliberately consider some very simple cases. Our experience is that for more complex theories the output quickly becomes more complicated. Although this is useful for automated computations, it is much less comprehensible by humans.

- Suppose that in the previous example one wants to impose a three-valued semantics. It is possible to do so by adding to Γ the restriction $\psi = \forall x(Q(x) \vee \neg Q(x))$, which is translated to $\overline{\psi} = \forall x(Q^+(x) \vee Q^-(x))$. Circumscribing $\overline{\Gamma} \cup \{\overline{\psi}\}$ yields

$$\forall x \; \{ \; [(Q^+(x) \wedge x \neq a \wedge x \neq b) \rightarrow \neg Q^-(x)] \; \wedge \; [(Q^-(x) \wedge x \neq a) \rightarrow \neg Q^+(x)] \; \},$$

  which has almost the same meaning as before, except that this time, the combination of this and $\overline{\psi}$ means that if $c \notin \{a, b\}$ then either $Q^+(c)$ or $Q^-(c)$ holds, but not both. It follows, then, that for such c, $Q(c)$ must have some classical value. Again, this corresponds to what one expects when k-minimizing $\Gamma \cup \{\psi\}$.

4 Using more than four values

In this section we extend to the multiple-valued case the results obtained in the previous sections. Essentially, we go through the same process. We start with a multiple-valued logic that *may not be paraconsistent*. Then we shift to a different semantics that is based on a more complex structure of truth values, and then show how to simulate the latter semantics by the original one. The outcome of this process is twofold:

1. A general approach for deriving a paraconsistent logic from a general multiple-valued logic.
2. A reduction of (preferential) paraconsistent reasoning in this derived logic to object-level reasoning in the original logic.

We thus obtain a general method for performing paraconsistent reasoning in a multi-valued logic. This extension may serve as an evidence for the robustness of the techniques proposed in this paper.

4.1 Lattice-valued reasoning

Definition 4.1 A *multiple-valued structure* for a language Σ is a triple $(\mathcal{L}, \mathcal{D}_\mathcal{L}, \mathcal{O}_\mathcal{L})$, where $\mathcal{L}$ is set of elements ("truth values"), $\mathcal{D}_\mathcal{L}$ is a nonempty proper subset of $\mathcal{L}$, and $\mathcal{O}_\mathcal{L}$ is a set of operations on $\mathcal{L}$ that correspond to the connectives in Σ.

In the sequel we shall assume that $\mathcal{L} = (L, \leq_L)$ is a complete lattice with a negation operator $\neg$,[16] and that $\mathcal{D}_\mathcal{L}$ is a filter in it, namely: it is a nonempty proper subset of L s.t. for every $x, y \in L$, $x \wedge y \in \mathcal{D}_\mathcal{L}$ iff $x \in \mathcal{D}_\mathcal{L}$ and $y \in \mathcal{D}_\mathcal{L}$. Sometimes we shall assume that $\mathcal{D}_\mathcal{L}$ is a *prime* filter in $\mathcal{L}$, i.e. that it is a filter in $\mathcal{L}$ s.t. for every $x, y \in L$, $x \vee y \in \mathcal{D}_\mathcal{L}$ iff $x \in \mathcal{D}_\mathcal{L}$ or $y \in \mathcal{D}_\mathcal{L}$.

The set $\mathcal{D}_\mathcal{L}$ consists of the *designated* values of $\mathcal{L}$, i.e., those that represent true assertions. By its definition it is obvious that $\mathcal{D}_\mathcal{L}$ is $\leq_L$-upwards closed, and so $\sup(L) \in \mathcal{D}_\mathcal{L}$ and $\inf(L) \notin \mathcal{D}_\mathcal{L}$.

[16] That is, for every $x, y \in L$, $x \leq_L y$ iff $\neg y \leq_L \neg x$, and for every $x \in L$, $\neg\neg x = x$. The requirement for a *complete* lattice is needed here for giving semantics to quantified formulae: for a structure with a domain E and a valuation ν, we let $\nu(\forall x \psi(x)) = \inf_{\leq_t}\{\nu(\psi(e)) \mid e \in E\}$; for all other purposes, it is sufficient to take distributive lattices with negation operators.

In what follows we further on assume that Σ is the classical propositional language where conjunctions correspond to the meet operation in $\mathcal{L}$, disjunctions corresponds to the join operation in $\mathcal{L}$, and negations correspond to the negation operator of $\mathcal{L}$.

Definition 4.2 Let $\mathcal{L}$ be a complete lattice, $\mathcal{D}_{\mathcal{L}}$ a prime filter in it, and Γ a set of formulae in Σ.

a) A (multiple-valued) *valuation* ν is a function that assigns an element of L to each atomic formula. A valuation is extended to complex formulae in the standard way.

b) A valuation ν *satisfies* a formula ψ if $\nu(\psi) \in \mathcal{D}_{\mathcal{L}}$.

c) A valuation ν is a *model* of Γ if it satisfies every formula in Γ. We shall continue to denote by $mod(\Gamma)$ the set of the models of Γ.

Definition 4.3 Let $\mathcal{L}$ be a complete lattice and $\mathcal{D}_{\mathcal{L}}$ a prime filter in it. For a set Γ of formulae and a formula ψ, denote $\Gamma \models^{\mathcal{L},\mathcal{D}_{\mathcal{L}}} \psi$ if every model of Γ is a model of ψ.

Note that classical logic is obtained from the above definitions by taking the two-valued lattice $(\{t, f\}, f <_L t)$ with $\mathcal{D}_{\mathcal{L}} = \{t\}$. Similarly, Kleene three-valued logic [21] is obtained by taking the three-valued lattice $\mathcal{L} = (\{t, f, \bot\}, \leq_L)$ with $\mathcal{D}_{\mathcal{L}} = \{t\}$. The connectives in $\mathcal{O}_{\mathcal{L}}$ correspond in this case to the lattice operations of a lattice in which $f <_L \bot <_L t$ together with a negation operation defined by: $\neg f = t, \neg t = f, \neg\bot = \bot$. Note also that both these logics are *not* paraconsistent.

4.2 Paraconsistent reasoning through bilattices

In order to add paraconsistent capabilities to the lattice-based logic under consideration, we use the same methodology as in the two-valued case. The basic idea is to consider a logic in which the truth values are not the lattice elements but are arbitrary pairs of lattice elements, rather than pairs of $\{0, 1\}$, as in the case of $\mathcal{FOUR}$.

Definition 4.4 [18] Let $\mathcal{L} = (L, \leq_L)$ be a complete lattice. The structure $\mathcal{L} \odot \mathcal{L} = (L \times L, \leq_t, \leq_k, \neg)$ consists of pairs of elements from L that are arranged in two lattice structures as follows:

- $(L \times L, \leq_t)$, where $(y_1, y_2) \leq_t (x_1, x_2)$ iff $y_1 \leq_L x_1$ and $y_2 \geq_L x_2$
- $(L \times L, \leq_k)$, where $(y_1, y_2) \leq_k (x_1, x_2)$ iff $y_1 \leq_L x_1$ and $y_2 \leq_L x_2$

The unary operation $\neg$ is defined on $L \times L$ by $\neg(x_1, x_2) = (x_2, x_1)$.

The structure that $\mathcal{L} \odot \mathcal{L}$ forms is called a *bilattice* [16, 18]. As in the four-valued case, a truth value $(x, y) \in \mathcal{L} \odot \mathcal{L}$ may intuitively be understood so that x represents the amount of evidence *for* an assertion, while y represents the amount of evidence *against* it. It is easy to verify that the $\leq_k$-minimal element of $\mathcal{L} \odot \mathcal{L}$ is $(\inf(L), \inf(L))$, the $\leq_k$-maximal one is $(\sup(L), \sup(L))$, the $\leq_t$-minimal element is $(\inf(L), \sup(L))$, and the $\leq_t$-maximal one is $(\sup(L), \inf(L))$.

Example 4.5 Belnap's four-valued lattice $\mathcal{FOUR}$, considered in the previous sections, is a particular case of the algebraic structures defined in 4.4, since $\mathcal{FOUR} = \mathcal{TWO} \odot \mathcal{TWO}$, where $\mathcal{TWO}$ is the two-valued classical lattice. For another example, consider the three-valued lattice $\mathcal{THREE} = (\{0, \frac{1}{2}, 1\},\ 0 < \frac{1}{2} < 1)$. Figure 2 contains a double-Hasse diagram of $\mathcal{THREE} \odot \mathcal{THREE}$.

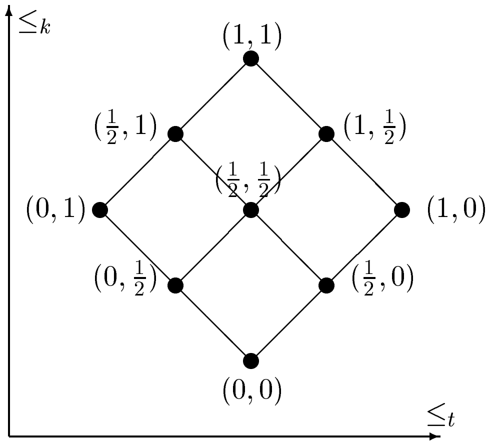

Figure 2: $\mathcal{THREE} \odot \mathcal{THREE}$

In what follows we shall continue to use the symbols $\vee, \wedge, \oplus$, and $\otimes$ for denoting, respectively, the $\leq_t$-join, $\leq_t$-meet, $\leq_k$-join, and the $\leq_k$-meet operations in $\mathcal{L} \odot \mathcal{L}$. By Definition 4.4 it follows that these operations are computed as in $\mathcal{FOUR}$, i.e., for every $x_1, x_2, y_1, y_2 \in L$,

$$(x_1, y_1) \vee (x_2, y_2) = (x_1 \vee x_2,\ y_1 \wedge y_2), \qquad (x_1, y_1) \wedge (x_2, y_2) = (x_1 \wedge x_2,\ y_1 \vee y_2),$$

$$(x_1, y_1) \oplus (x_2, y_2) = (x_1 \vee x_2,\ y_1 \vee y_2), \qquad (x_1, y_1) \otimes (x_2, y_2) = (x_1 \wedge x_2,\ y_1 \wedge y_2).$$

As noted in Example 4.5, Definition 4.4 is a natural extension of Belnap's four-valued structure. The notion of the designated values in $\mathcal{FOUR}$ can also be generalized in a natural way in $\mathcal{L} \odot \mathcal{L}$:

Definition 4.6 [2] Let $\mathcal{L} \odot \mathcal{L}$ be the bilattice defined in 4.4.

a) A *bifilter* $\mathcal{D}$ of $\mathcal{L} \odot \mathcal{L}$ is a nonempty proper subset of $L \times L$, such that:
   (i) $x \wedge y \in \mathcal{D}$ iff $x \in \mathcal{D}$ and $y \in \mathcal{D}$, (ii) $x \otimes y \in \mathcal{D}$ iff $x \in \mathcal{D}$ and $y \in \mathcal{D}$.

b) A bifilter $\mathcal{D}$ is called *prime*, if it also satisfies the following conditions:
   (i) $x \vee y \in \mathcal{D}$ iff $x \in \mathcal{D}$ or $y \in \mathcal{D}$, (ii) $x \oplus y \in \mathcal{D}$ iff $x \in \mathcal{D}$ or $y \in \mathcal{D}$.

Given a bilattice of the form $\mathcal{L} \odot \mathcal{L}$, we fix some prime bifilter $\mathcal{D}$ in $\mathcal{L} \times \mathcal{L}$. This set consists of the *designated elements* of $\mathcal{L} \odot \mathcal{L}$. It is easy to verify that a prime bifilter of $\mathcal{L} \odot \mathcal{L}$ is upwards closed w.r.t. both partial orders of $\mathcal{L} \odot \mathcal{L}$, thus it contains the $\leq_t$- and the $\leq_k$-largest element and does not contain the $\leq_t$- and the $\leq_k$-least one.

As in the lattice-valued case,(prime) bifilters are used for defining validity of formulae: a valuation ν on $\mathcal{L} \times \mathcal{L}$ is a *model* of a set Γ of formulae if $\nu(\psi) \in \mathcal{D}$ for every $\psi \in \Gamma$.

Example 4.7 The set $\mathcal{D} = \{t, \top\}$ of the designated elements in $\mathcal{FOUR}$ is indeed a prime bifilter in $\mathcal{FOUR}$ (and, moreover, it is the only prime bifilter in this bilattice). In $\mathcal{THREE} \odot \mathcal{THREE}$ there are two prime bifilters:

- $\mathcal{D}_1 = \{(1,x) \mid x \in \{0, \frac{1}{2}, 1\}\} = \{(x_1,x_2) \mid (x_1,x_2) \geq_k (1,0)\}$, and
- $\mathcal{D}_2 = \{(x_1,x_2) \mid x_1 \geq \frac{1}{2},\ x_2 \in \{0, \frac{1}{2}, 1\}\} = \{(x_1,x_2) \mid (x_1,x_2) \geq_k (\frac{1}{2},0)\}$.

Proposition 4.8 [5]

a) $\mathcal{D}$ is a bifilter in $\mathcal{L} \odot \mathcal{L}$ iff $\mathcal{D} = \mathcal{D}_{\mathcal{L}} \times \mathcal{L}$, where $\mathcal{D}_{\mathcal{L}}$ is a filter in $\mathcal{L}$.

b) $\mathcal{D}$ is a prime bifilter in $\mathcal{L} \odot \mathcal{L}$ iff $\mathcal{D} = \mathcal{D}_{\mathcal{L}} \times \mathcal{L}$, where $\mathcal{D}_{\mathcal{L}}$ is a prime filter in $\mathcal{L}$.

Corollary 4.9 [5] Let $x_0 \in L$, $x_0 \neq \inf(L)$. Denote: $\mathcal{D}(x_0) = \{(y_1,y_2) \mid y_1 \geq_L x_0,\ y_2 \in L\}$, and $\mathcal{D}_{\mathcal{L}}(x_0) = \{y \in L \mid y \geq_L x_0\}$. Then:

a) $\mathcal{D}(x_0)$ is a prime bifilter in $\mathcal{L} \odot \mathcal{L}$ iff $\mathcal{D}_{\mathcal{L}}(x_0)$ is a prime filter in $\mathcal{L}$.

b) $\mathcal{D}(\sup(L))$ is a prime bifilter in $\mathcal{L} \odot \mathcal{L}$ iff $\sup(L)$ is *join irreducible* (i.e., iff $x \vee_{\mathcal{L}} y = \sup(L)$ implies that $x = \sup(L)$ or $y = \sup(L)$).

c) If the condition of item (b) is met, then $\mathcal{D}(\sup(L))$ is least (w.r.t. set inclusion) among the (prime) bifilters of $\mathcal{L} \odot \mathcal{L}$.

Given a billatice $\mathcal{L} \odot \mathcal{L}$ and a prime bifilter $\mathcal{D}$ in it, one may define a corresponding consequence relation in a similar way to the lattice-valued case:

Definition 4.10 $\Gamma \models^{\mathcal{L}\odot\mathcal{L},\mathcal{D}} \psi$ if every $\mathcal{L} \times \mathcal{L}$-valued model of Γ is a $\mathcal{L} \times \mathcal{L}$-valued model of ψ.

Proposition 4.11 $\models^{\mathcal{L}\odot\mathcal{L},\mathcal{D}}$ is paraconsistent.

Proof: It is easy to verify that $\models^{\mathcal{L}\odot\mathcal{L},\mathcal{D}}$ does not allow trivial reasoning from inconsistent theories. Indeed, $p, \neg p \not\models^{\mathcal{L}\odot\mathcal{L},\mathcal{D}} q$. A counter-model assigns $(\sup(L), \sup(L))$ to p and $(\inf(L), \sup(L))$ to q. $\square$

4.3 Simulating bilattice-valued reasoning

In the rest of this paper we fix some lattice $\mathcal{L} = (L, \leq_L)$ with a prime filter $\mathcal{D}_{\mathcal{L}}$, and denote by $\mathcal{D}$ the corresponding prime bifilter of the form $\mathcal{D}_{\mathcal{L}} \times \mathcal{L}$ in $\mathcal{L} \odot \mathcal{L}$. We shall show that by using the same method as that of Section 3 for the four-valued case, it is now possible to have analogous results for every structure of the form $\mathcal{L} \odot \mathcal{L}$ with a set of designated values $\mathcal{D}$. Again, we start by standard bilattice-valued reasoning and then, in the next section, consider the preferential case.

For every formula ψ in Σ we can obtain its "split form", $\overline{\psi}$ in exactly the same way as in the four-valued case. Also, since every valuation ν on $\mathcal{L} \odot \mathcal{L}$ can be represented by a pair (ν_1, ν_2) of L-valued components, then $\overline{\nu}$ is an L-valued valuation defined (just as in the four-valued case) by $\overline{\nu}(p^+) = \nu_1(p)$ and $\overline{\nu}(p^-) = \nu_2(p)$. Again, whenever it is more convenient, we shall use the more detailed notation $(\vec{p}^{\,+} : \nu_1 \,;\, \vec{p}^{\,-} : \nu_2)$ instead of just $\overline{\nu}$.

Using the notations above we can now generalize Proposition 2.4 to the case of $\mathcal{L} \odot \mathcal{L}$ as follows:

Proposition 4.12 Let $\mathcal{L}$ be a de Morgan lattice,[17] and let $\nu = (\nu_1, \nu_2)$ be a valuation on $\mathcal{L} \odot \mathcal{L}$. Then $\nu(\psi) = (\overline{\nu}(\overline{\psi}), \neg\,\mathtt{rev}\,(\overline{\nu})(\overline{\psi}))$.

The proof of Proposition 4.12 is identical to that of Proposition 2.4, using $\mathcal{L} \odot \mathcal{L}$ instead of $\mathcal{FOUR}$.

In the rest of the paper we suppose, then, that $\mathcal{L}$ is a de Morgan lattice.[18] Under this assumption, the next two corollaries immediately follow from Proposition 4.12.

Corollary 4.13 Let $\mathcal{D}_{\mathcal{L}}$ be a prime filter in $\mathcal{L}$, and let $\mathcal{D} = \mathcal{D}_{\mathcal{L}} \times \mathcal{L}$ be the set of the designated elements in $\mathcal{L} \odot \mathcal{L}$.[19] For every valuation ν on $\mathcal{L} \odot \mathcal{L}$ and a formula ψ in Σ, $\nu(\psi) \in \mathcal{D}$ iff $\overline{\nu}(\overline{\psi}) \in \mathcal{D}_{\mathcal{L}}$.

Proof: $\nu(\psi)$ is designated iff $\nu_1(\psi) \in \mathcal{D}_{\mathcal{L}}$, iff (Proposition 4.12) $\overline{\nu}(\overline{\psi}) \in \mathcal{D}_{\mathcal{L}}$. □

In particular, since by Proposition 4.8(b) *every* prime bifilter in $\mathcal{L} \times \mathcal{L}$ is of the form $\mathcal{D}_{\mathcal{L}} \times \mathcal{L}$, where $\mathcal{D}_{\mathcal{L}}$ is a prime filter in $\mathcal{L}$, we have that whenever $\nu(\psi)$ is designated in $\mathcal{L} \times \mathcal{L}$, there is a prime filter in $\mathcal{L}$ with respect to which $\overline{\nu}(\overline{\psi})$ is designated in $\mathcal{L}$, and vice-versa.

Corollary 4.14 Suppose that $\sup(L)$ is join irreducible in $\mathcal{L}$, and let $\mathcal{D} = \mathcal{D}(\sup(L))$ be the set of the designated elements in $\mathcal{L} \odot \mathcal{L}$.[20] For every valuation ν in $\mathcal{L} \odot \mathcal{L}$ and a formula ψ in Σ, $\nu(\psi) \in \mathcal{D}$ iff $\overline{\nu}(\overline{\psi}) = \sup(L)$.

Proof: $\nu(\psi)$ is designated iff $\nu_1(\psi) = \sup(L)$, iff (Proposition 4.12) $\overline{\nu}(\overline{\psi}) = \sup(L)$. □

We can now extend Theorem 2.8, and show how to simulate reasoning in $\mathcal{L} \odot \mathcal{L}$ by object level reasoning in $\mathcal{L}$.

Theorem 4.15 $\Gamma \models^{\mathcal{L} \odot \mathcal{L}, \mathcal{D}} \psi$ iff $\overline{\Gamma} \models^{\mathcal{L}, \mathcal{D}_{\mathcal{L}}} \overline{\psi}$.

Proof: Follows from Corollary 4.13. □

As in the four-valued case, since in the language without negations, $\overline{\Gamma}$ is obtained just by renaming the atomic formulae that appear in Γ, the following corollary immediately follows from Theorem 4.15 (cf. Corollary 2.9):

Corollary 4.16 In positive logic, $\Gamma \models^{\mathcal{L} \odot \mathcal{L}, \mathcal{D}} \psi$ iff $\Gamma \models^{\mathcal{L}, \mathcal{D}_{\mathcal{L}}} \psi$.

---

[17] That is, for every $x, y \in L$, $\neg(x \vee y) = \neg x \wedge \neg y$, and $\neg(x \wedge y) = \neg x \vee \neg y$.

[18] It is interesting to note that $\mathcal{L} \odot \mathcal{L}$ is always a de Morgan bilattice. Indeed, $\neg[(x_1, x_2) \vee (y_1, y_2)] = \neg(x_1 \vee y_1, x_2 \wedge y_2) = (x_2 \wedge y_2, x_1 \vee y_1) = (x_2, x_1) \wedge (y_2, y_1) = \neg(x_1, x_2) \wedge \neg(y_1, y_2)$.

[19] By Proposition 4.8, $\mathcal{D}$ is indeed a prime bifilter in $\mathcal{L} \odot \mathcal{L}$.

[20] By Corollary 4.9, $\mathcal{D}$ is indeed a prime bifilter in $\mathcal{L} \odot \mathcal{L}$.

4.4 Simulating bilattice-valued preferential reasoning

We turn now to the preferential case. Again, by using the same methods as those of Section 3 for the four-valued case, it is possible to define circumscriptive formulae for expressing multiple-valued preferential reasoning w.r.t. a representable preference order $\leq$. This implies, in particular, that once again we are able to reduce "meta-reasoning" (this time, in the bilattice) to object level reasoning (in the lattice) by axiomatizing the preferred models of a given theory in the bilattice through formula circumscription in the lattice.

Suppose, then, that $\leq$ is some preferential order among the valuations into $\mathcal{L}\times\mathcal{L}$. Note that the notion of a representation of $\leq$ by a formula $\Psi_{\leq}$, defined in 3.12, can be directly extended to the bilattice-valued case. A corresponding consequence relation can also be defined by extending the definition of the four-valued case:

Notation 4.17 $\Gamma \models_{\leq}^{\mathcal{L}\odot\mathcal{L},\mathcal{D}} \psi$ iff every $\leq$-most preferred $\mathcal{L}\odot\mathcal{L}$-valued model of Γ is a $\mathcal{L}\odot\mathcal{L}$-valued model of ψ.[21]

Now, for constructing the circumscribing formulae in the L-valued logic we need to have an implication connective on L that behaves as material implication in classical logic. This is what we define next.

Definition 4.18 Let $\mathcal{L}=(L,\leq_L)$ be a lattice and $\mathcal{D}_{\mathcal{L}}$ a (prime) filter in it. For every $x,y\in L$ define:

a) $x\to y = \sup(L)$ if either $x\notin\mathcal{D}_{\mathcal{L}}$ or $y\in\mathcal{D}_{\mathcal{L}}$, otherwise $x\to y = \inf(L)$.

b) $x\preceq y = \sup(L)$ if $x\leq_L y$, otherwise $x\preceq y = \inf(L)$.

c) $x\prec y = \sup(L)$ if $x<_L y$, otherwise $x\prec y = \inf(L)$. [22]

Our next result extends Theorem 3.16:

Theorem 4.19 Let Γ be a set of formula and ψ a formula in Σ. Let $\mathsf{Circ}_{\leq}$ be the formula given in Proposition 3.13 for a preferential relation $\leq$.[23] Then $\Gamma \models_{\leq}^{\mathcal{L}\odot\mathcal{L},\mathcal{D}} \psi$ iff $\overline{\Gamma}\cup\mathsf{Circ}_{\leq} \models^{\mathcal{L},\mathcal{D}_{\mathcal{L}}} \overline{\psi}$.

Proof: As in the proof of Theorem 3.16, the claim follows from the fact that ν is a $\leq$-most preferred model of ψ iff $\overline{\nu}$ satisfies $\overline{\psi}$ (this is true by Corollary 4.13) and the fact that $\overline{\nu}$ also satisfies $\mathsf{Circ}_{\leq}$ (the proof of Proposition 3.13 may be used in the present case as well for showing the latter fact). $\square$

Clearly, Theorem 4.19 may be applied only in cases that the preferential relations under consideration are *representable* (in the sense of Definition 3.12). Next we show that all the preferential relations that have been considered in Section 3 can be generalized to (bi)lattice-valued preferential relations that are also representable by circumscriptive formulae.

[21] Again, this is the same relation as the one defined in 3.3, together with an explicit indication that the underlying semantics is based on $\mathcal{L}\odot\mathcal{L}$ and $\mathcal{D}$.

[22] Note that this is a generalization of the definition of the operators with the same notations, given in Notation 3.17. In particular, when L is the two-valued lattice, $\to$ and $\preceq$ are the same as the classical implication.

[23] Where the relevant connectives are interpreted by Definition 4.18.

Representing preference of minimal information

As in Proposition 3.19, the set of the k-minimal models of ψ in $\mathcal{L} \odot \mathcal{L}$ can be represented by $\mathsf{Circ}_{\leq_k}$, using the generalized interpretations for the relevant connectives (see Definition 4.18):

Proposition 4.20 A valuation ν in $\mathcal{L} \odot \mathcal{L}$ is a k-minimal model of ψ iff $\overline{\nu}(\overline{\psi}) \in \mathcal{D}_{\mathcal{L}}$ and $\overline{\nu}(\mathsf{Circ}_{\leq_k}) = \sup(L)$.

Proof: The same proof as that of Proposition 3.13 holds in this case as well (where every reference to Corollary 2.7 should be replaced by a reference to Corollary 4.13). Thus, ν is a k-minimal model of ψ iff $\overline{\nu}(\overline{\psi}) \in \mathcal{D}_{\mathcal{L}}$ and $\overline{\nu}(\mathsf{Circ}_{\leq_k}) \in \mathcal{D}_{\mathcal{L}}$, where

$$\mathsf{Circ}_{\leq_k}(\vec{p}^{\,\pm}) \;=\; \forall(\vec{q}^{\,\pm})\,\{\,\overline{\psi}(\vec{q}^{\,\pm}) \;\to\; (\,\Psi_{\leq_k}(\vec{q}^{\,\pm},\vec{p}^{\,\pm}) \to \Psi_{\leq_k}(\vec{p}^{\,\pm},\vec{q}^{\,\pm})\,)\,\},$$

and

$$\Psi_{\leq_k}(\vec{p}^{\,\pm},\vec{q}^{\,\pm}) \;=\; \bigwedge_{i=1}^{n} ((p_i^+ \preceq q_i^+) \wedge (p_i^- \preceq q_i^-)).$$

Note also that $\mathsf{Circ}_{\leq_k}$ is defined only by operators onto $\{\inf(L), \sup(L)\}$. Now , since all the operators in Σ are closed w.r.t. this set of values, and since $\sup(L) \in \mathcal{D}_{\mathcal{L}}$, we have that $\overline{\nu}(\mathsf{Circ}_{\leq_k}) \in \mathcal{D}_{\mathcal{L}}$ iff $\overline{\nu}(\mathsf{Circ}_{\leq_k}) = \sup(L)$.

It remains to show that $\leq_k$ is represented by $\Psi_{\leq_k}$. Indeed, by the way the partial order $\leq_k$ is defined on $L \times L$ (see Definition 4.4) and by Definition 4.18, we have that

$$\begin{aligned}
\nu \leq_k \mu \iff& \forall 1 \leq i \leq n \;\, \nu(p_i) \leq_k \mu(p_i) \iff \forall 1 \leq i \leq n \;\, \nu_1(p_i) \leq_L \mu_1(p_i) \text{ and } \nu_2(p_i) \leq_L \mu_2(p_i) \\
\iff& (\vec{p}^{\,\pm}\!:\!\nu,\; \vec{q}^{\,\pm}\!:\!\mu) \text{ satisfies } \textstyle\bigwedge_{i=1}^{n} ((p_i^+ \preceq q_i^+) \wedge (p_i^- \preceq q_i^-)) \\
\iff& (\vec{p}^{\,\pm}\!:\!\nu,\; \vec{q}^{\,\pm}\!:\!\mu) \text{ satisfies } \Psi_{\leq_k}(\vec{p}^{\,\pm},\vec{q}^{\,\pm}).
\end{aligned}$$

□

Representing preference of most consistent interpretations

Let $\mathcal{I}_\top = \{x \in \mathcal{L} \times \mathcal{L} \mid x \in \mathcal{D}, \neg x \in \mathcal{D}\}$ be the set of the *inconsistent* values in $\mathcal{L} \odot \mathcal{L}$. A valuation ν_1 is (strictly) *more consistent* than a valuation ν_2 if the set of atoms p_i s.t. $\nu_1(p_i) \in \mathcal{I}_\top$ is (strictly) subsumed in the set of atoms p_j s.t. $\nu_2(p_j) \in \mathcal{I}_\top$. A valuation $\nu \in \mathit{mod}(\psi)$ is a *most consistent model* of ψ [4, 5], if there is no other model of ψ that is strictly more consistent than ν.

By a proof that is similar to that of Proposition 4.20 one can show that the set of the most consistent models of ψ can be represented by $\mathsf{Circ}_{\leq_{\{\top\}}}$: a valuation ν in $\mathcal{L} \odot \mathcal{L}$ is a most consistent model of ψ iff $\overline{\nu}(\overline{\psi}) \in \mathcal{D}_{\mathcal{L}}$ and $\overline{\nu}(\mathsf{Circ}_{\leq_{\{\top\}}}) = \sup(L)$.

Representing preference of most classical interpretations

Let $\mathcal{I}_\top$ be the set of the inconsistent elements in $\mathcal{L} \odot \mathcal{L}$ as in the previous case, and let $\mathcal{I}_\perp = \{x \in \mathcal{L} \times \mathcal{L} \mid x \notin \mathcal{D}, \neg x \notin \mathcal{D}\}$ be the set of the *incomplete values* in $\mathcal{L} \odot \mathcal{L}$. A valuation ν_1 is (strictly) *more classical* than a valuation ν_2 if the set of atoms p_i s.t. $\nu_1(p_i) \in \mathcal{I}_\top \cup \mathcal{I}_\perp$ is (strictly) subsumed in the set of atoms p_j s.t. $\nu_2(p_j) \in \mathcal{I}_\top \cup \mathcal{I}_\perp$. A valuation $\nu \in \mathit{mod}(\psi)$ is a *most classical model* of ψ [4, 5], if there is no other model of ψ that is strictly more classical than ν.

Again, a similar proof as that of Proposition 4.20 shows that the set of the most classical models of ψ can be represented by $\mathsf{Circ}_{\leq\{\top,\bot\}}$: a valuation ν in $\mathcal{L} \odot \mathcal{L}$ is a most classical model of ψ iff $\overline{\nu}(\overline{\psi}) \in \mathcal{D}_{\mathcal{L}}$ and $\overline{\nu}(\mathsf{Circ}_{\leq\{\top,\bot\}}) = \sup(L)$.

Particular cases in which the representations above may be used are the bilattice-based logics introduced in [2, 3, 4, 5], and the annotated logic [35] RI, introduced in [22, 23], provided that the underlying many-valued structure is of the form $\mathcal{L} \odot \mathcal{L}$.

5 Summary and conclusion

A well-known way of formalizing paraconsistent reasoning is in terms of de Morgan algebras, with a certain four-element algebra playing a pivotal role analogous to that of the two-element Boolean algebra in its class. To formalize reasoning that is simultaneously paraconsistent and non-monotonic, Belnap [8, 9] and Ginsberg [18, 19] have elaborated de Morgan algebras into bilattices. In [4] it is shown that the four-element bilattice $\mathcal{FOUR}$ again plays a pivotal role. In this paper we followed up this work, essentially motivated by computational considerations. We have shown that questions of consequence in these structures can be reduced to ones of classical consequence, by means of polynomial translations that essentially serves to separate negated atoms from affirmed ones. Moreover, these translations can be incorporated together with some appropriate circumscriptive axioms to capture the notion of minimality and for representing preferential reasoning [34]. This method also touches upon several additional aspects:

1. It shows that two-valued reasoning may be useful for simulating inference procedures in the context of many-valued semantics.

2. This approach demonstrates the usefulness of circumscription not only as a general method for non-monotonic inferential reasoning, but also as an appealing technique for implementing paraconsistent reasoning.

3. This is another evidence for the fact that in many cases concepts that are defined in a "meta-language" (such as preference criteria, etc.) can be expressed in the language itself (using, e.g., higher-order formulae).

Note that although we have proposed our technique for propositional logic, it can be easily applied to the predicate case as well. Moreover, as shown in Section 4, our approach can be extended to many-valued (lattice-based) logics. These observations, together with item (3) above, imply that such techniques allow a potentially wide area for practical implementations. For instance, as we have shown above, preferential multiple-valued reasoning can be incorporated with practical applications for automated reasoning and theorem proving.

Acknowledgement

We thank the anonymous referees for many useful comments.

References

[1] A.R.Anderson, N.D.Belnap. *Entailment* Vol.1, Princton University Press, 1975.

[2] O.Arieli, A.Avron. *Reasoning with logical bilattices.* Journal of Logic, Language, and Information 5(1), pp. 25–63, 1996.

[3] O.Arieli, A.Avron. *The logical role of the four-valued bilattice.* Proc. 13th IEEE Conf. on Logic in Computer Science (LICS'98), pp. 218–226, IEEE Press, 1998.

[4] O.Arieli, A.Avron. *The value of the four values.* Artificial Intelligence 102(1), pp. 97–141, 1998.

[5] O.Arieli, A.Avron. *Bilattices and paraconsistency.* Frontiers of Paraconsistent Logic (D. Batens, C. Mortensen, G. Priest, J. Van Bendegem, editors), pp. 11–27, Studies in Logic and Computation 8, Research Studies Press, 2000.

[6] O.Arieli, M.Denecker. *Modeling paraconsistent reasoning by classical logic* Proc. 2nd Int. Symp. on Foundations of Information and Knowledge Systems (FoIKS'02). Lecture Notes in Computer Science 2284 (T.Eiter, K.D.Schewe, editors), pp. 1–14, Springer, 2002.

[7] A.Avron. *Simple consequence relations.* Information and Computation 92, pp. 105–139, 1991.

[8] N.D.Belnap. *A useful four-valued logic.* Modern Uses of Multiple-Valued Logic (G.Epstein, J.M.Dunn, editors), Reidel Publishing Company, Boston, pp. 7–37, 1977.

[9] N.D.Belnap. *How a computer should think.* Contemporary Aspects of Philosophy (G.Ryle, editor), pp. 30–56, Oriel Press, 1977.

[10] A.Bialynicki-Birula. *Remarks on quasi-boolean algebras.* Bull. Acad. Polonaise des Sciences Cl. III, Vol. V No. 6, pp. 615–619, 1957.

[11] A.Bialynicki-Birula, H.Rasiowa. *On the representation of quasi-boolean algebras*. Bull. Acad. Polonaise des Sciences Cl. III, Vol. V, No. 3, pp. 259–261, 1957.

[12] N.C.A.da-Costa. *On the theory of inconsistent formal systems.* Notre Damm Journal of Formal Logic 15, pp. 497–510, 1974.

[13] P.Doherty, W.Lukaszewicz, A.Szalas. *Computing circumscription revisited: Preliminary report.* Proc. 14th Int. Joint Conf. on Artificial Intelligence (IJCAI'95), pp. 1502–1508, 1995.

[14] P.Doherty, W.Lukaszewicz, A.Szalas. *Computing circumscription revisited: A reduction algorithm.* Journal of Automated Reasoning 18, pp. 297–334, 1997.

[15] M.Fitting. *Kripke-Kleene semantics for logic programs.* Journal of Logic Programming 2, pp. 295–312, 1985.

[16] M.Fitting. *Kleene's logic, generalized.* Journal of Logic and Computation 1, pp. 797–810, 1990.

[17] D.Gabbay, O.Rodrigues, A.Russo. *Revision by translation.* Information, Uncertainty, and Fusion (B.Bouchon-Meunier, R.R.Yager, L.Zadeh, editors), pp. 3–32, Kluwer Academic Publishers, 2000.

[18] M.L.Ginsberg. *Multi-valued logics: A uniform approach to reasoning in AI.* Computer Intelligence 4, pp. 256–316, 1988.

[19] M.L.Ginsberg. *A circumscriptive theorem prover.* Artificial Intelligence 29, pp. 209–230, 1989.

[20] J.A.Kalman. *Lattices with involution.* Trans. of the American Math. Soc. 87, pp. 485–491, 1958.

[21] S.C.Kleene. *Introduction to metamathematics.* Van Nostrand, 1950.

[22] M.Kifer, E.L.Lozinskii. *RI: A logic for reasoning with inconsistency.* Proc. IEEE Conf. on Logic in Computer Science (LICS'89), pp. 253–262, IEEE Press, 1989.

[23] M.Kifer, E.L.Lozinskii. *A logic for reasoning with inconsistency.* Journal of Automated Reasoning 9(2), pp. 179–215, 1992.

[24] S.Kraus, D.Lehmann, M.Magidor. *Nonmonotonic reasoning, preferential models and cumulative logics.* Artificial Intelligence 44(1–2) pp. 167–207, 1990.

[25] D.Lehmann, M.Magidor. *What does a conditional knowledge base entail?* Artificial Intelligence 55, pp. 1–60, 1992.

[26] D.Makinson. *General theory of cumulative inference.* Non-Monotonic Reasoning (M.Reinfrank, editor), Lecture Notes in Artificial Intelligence 346, pp. 1–18, Springer, 1989.

[27] D.Makinson. *General patterns in nonmonotonic reasoning.* Handbook of Logic in Artificial Intelligence and Logic Programming 3 (D.Gabbay, C.Hogger, J.Robinson, editors) pp. 35–110, Oxford Science Pub., 1994.

[28] J.McCarthy. *Circumscription – A form of non monotonic reasoning.* Artificial Intelligence 13(1–2), pp. 27–39, 1980.

[29] J. McCarthy. *Applications of circumscription to formalizing common-Sense knowledge.* Artifical Intelligence 28, pp. 89–116, 1986.

[30] H.J.Ohlbach, *SCAN – Elimination of predicate quantifiers.* Proc. of the 13th conf. on automated deductiona (CADE'96), (M.A.McRobbie, J.Slaney, editors), Lecture Notes in Artificial Intelligence 1104, pp. 161–165, Springer, 1996.

[31] G.Priest. *Reasoning about truth.* Artificial Intelligence 39, pp. 231–244, 1989.

[32] G.Priest. *Minimally Inconsistent LP.* Studia Logica 50, pp. 321–331, 1991.

[33] T.Przymusinski. *An algorithm to compute circumscription.* Artificial Intelligence 38, pp. 49–73, 1991.

[34] Y.Shoham. *Reasoning about change.* MIT Press, 1988.

[35] V.S.Subrahmanian. *Mechanical proof procedures for many valued lattice-based logic programming.* Journal of Non-Classical Logic 7, pp. 7–41, 1990.

[36] A.Tarski. *Introduction to logic.* Oxford University Press, N.Y., 1941.

[37] A.Van Gelder, K.A.Ross and J.S.Schlipf. *The well-founded semantics for general logic programs.* Journal of the ACM 38(3), pp. 620–650, 1991.

Appendix A. Proof of Proposition 2.4

Proposition 2.4: Let $\nu = (\nu_1, \nu_2)$. Then $\nu(\psi) = (\overline{\nu}(\overline{\psi}),\ \neg\texttt{rev}(\overline{\nu})(\overline{\psi}))$.

Proof: Recall that $\overline{\nu} = (\vec{p}^{\,+} : \nu_1 \,;\, \vec{p}^{\,-} : \nu_2)$ and that $\texttt{rev}(\overline{\nu}) = (\vec{p}^{\,+} : \neg\nu_2 \,;\, \vec{p}^{\,-} : \neg\nu_1)$.

We first show two lemmas:

Lemma 2.4-A: $\texttt{rev}(\overline{\nu})(\neg\overline{\psi}) = \overline{\nu}(\overline{\neg\psi})$.

Lemma 2.4-B: $\overline{\nu}(\neg\overline{\psi}) = \texttt{rev}(\overline{\nu})(\overline{\neg\psi})$.

Proof of Lemmas 2.4-A and 2.4-B: Note first that both lemmas are equivalent. Indeed, since $\texttt{rev}(\texttt{rev}(\overline{\nu})) = \overline{\nu}$, by replacing $\overline{\nu}$ by $\texttt{rev}(\overline{\nu})$ in Lemma 2.4-A we obtain Lemma 2.4-B and vice versa.

The proof is by the following induction on the structure of ψ:

- $\psi = p$:

  $\texttt{rev}(\overline{\nu})(\neg\overline{\psi}) = \texttt{rev}(\overline{\nu})(\neg p^+) = \neg(\vec{p}^{\,+} : \neg\nu_2 \,;\, \vec{p}^{\,-} : \neg\nu_1)(p^+) = \neg\neg\nu_2(p) = \nu_2(p)$.
  On the other hand,
  $\overline{\nu}(\overline{\neg\psi}) = \overline{\nu}(\overline{\neg p}) = \overline{\nu}(\neg\neg p^-) = (\vec{p}^{\,+} : \nu_1 \,;\, \vec{p}^{\,-} : \nu_2)(p^-) = \nu_2(p)$.

- $\psi = \neg\phi$:

  $\texttt{rev}(\overline{\nu})(\neg\overline{\psi}) = \texttt{rev}(\overline{\nu})(\neg\overline{\neg\phi}) = \neg\texttt{rev}(\overline{\nu})(\overline{\neg\phi}) =$ (by induction hypothesis) $=$
  $\neg\overline{\nu}(\neg\overline{\phi}) = \overline{\nu}(\neg\neg\overline{\phi}) = \overline{\nu}(\overline{\phi}) = \overline{\nu}(\overline{\neg\neg\phi}) = \overline{\nu}(\overline{\neg\psi})$.

- $\psi = \phi_1 \vee \phi_2$:

  By the definition of splitting transformation and by Definition 2.1, it is obvious that $\overline{\phi_1 \vee \phi_2} = \overline{\phi_1} \vee \overline{\phi_2}$ and $\overline{\phi_1 \wedge \phi_2} = \overline{\phi_1} \wedge \overline{\phi_2}$. Thus:

  $\texttt{rev}(\overline{\nu})(\neg\overline{\psi}) = \neg\texttt{rev}(\overline{\nu})(\overline{\psi}) = \neg\texttt{rev}(\overline{\nu})(\overline{\phi_1 \vee \phi_2}) = \neg[\texttt{rev}(\overline{\nu})(\overline{\phi_1}) \vee \texttt{rev}(\overline{\nu})(\overline{\phi_2})] =$
  (since $\mathcal{L}$ is a de Morgan lattice) $= \neg\texttt{rev}(\overline{\nu})(\overline{\phi_1}) \wedge \neg\texttt{rev}(\overline{\nu})(\overline{\phi_2}) =$
  $\texttt{rev}(\overline{\nu})(\neg\overline{\phi_1}) \wedge \texttt{rev}(\overline{\nu})(\neg\overline{\phi_2}) =$ (by induction hypothesis) $= \overline{\nu}(\overline{\neg\phi_1}) \wedge \overline{\nu}(\overline{\neg\phi_2}) =$
  $\overline{\nu}(\overline{\neg\phi_1} \wedge \overline{\neg\phi_2}) = \overline{\nu}(\overline{\neg\phi_1 \wedge \neg\phi_2}) =$ (de Morgan law again) $= \overline{\nu}(\overline{\neg(\phi_1 \vee \phi_2)}) = \overline{\nu}(\overline{\neg\psi})$.

- The case in which $\psi = \phi_1 \wedge \phi_2$ is analogue to the latter case.

Now we are ready to show the equation of Proposition 2.4. Again, we show it by an induction on the structure of ψ.

- $\psi = p$:

$$(\overline{\nu}(\overline{\psi}),\ \neg\mathtt{rev}\,(\overline{\nu})(\overline{\psi})) = (\overline{\nu}(p^+),\ \neg\mathtt{rev}\,(\overline{\nu})(p^+)) =$$
$$((\vec{p}^{\,+}\!:\!\nu_1\ ;\ \vec{p}^{\,-}\!:\!\nu_2)(p^+),\ \neg(\vec{p}^{\,+}\!:\!\neg\nu_2\ ;\ \vec{p}^{\,-}\!:\!\neg\nu_1)(p^+)) =$$
$$(\nu_1(p),\ \neg\neg\nu_2(p)) = (\nu_1(p),\ \nu_2(p)) = \nu(\psi).$$

- $\psi = \neg\phi$:

$$\nu(\psi) = \nu(\neg\phi) = \neg\nu(\phi) = \text{(by induction hypothesis)} = \neg(\ \overline{\nu}(\overline{\phi}),\ \neg\mathtt{rev}\,(\overline{\nu})(\overline{\phi})\ ) =$$
$$(\neg\mathtt{rev}\,(\overline{\nu})(\overline{\phi}),\ (\overline{\nu}(\overline{\phi}))) = (\mathtt{rev}\,(\overline{\nu})(\neg\overline{\phi}),\ \neg(\overline{\nu}(\neg\overline{\phi}))) = \text{(by Lemmas 2.4-A and 2.4-B)}$$
$$(\overline{\nu}(\overline{\neg\phi}),\ \neg\mathtt{rev}\,(\overline{\nu})(\overline{\neg\phi})) = (\overline{\nu}(\overline{\psi}),\ \neg\mathtt{rev}\,(\overline{\nu})(\overline{\psi})).$$

- $\psi = \phi_1 \vee \phi_2$:

$$\nu(\psi) = \nu(\phi_1 \vee \phi_2) = \nu(\phi_1) \vee \nu(\phi_2) = \text{(by induction hypothesis)} =$$
$$(\overline{\nu}(\overline{\phi_1}),\ \neg\mathtt{rev}\,(\overline{\nu})(\overline{\phi_1}))\ \vee\ (\overline{\nu}(\overline{\phi_2}),\ \neg\mathtt{rev}\,(\overline{\nu})(\overline{\phi_2})) = \text{(by the definition on } \vee) =$$
$$(\overline{\nu}(\overline{\phi_1}) \vee \overline{\nu}(\overline{\phi_2}),\ \neg\mathtt{rev}\,(\overline{\nu})(\overline{\phi_1}) \wedge \neg\mathtt{rev}\,(\overline{\nu})(\overline{\phi_2})) = \text{(by de Morgan law)} =$$
$$(\overline{\nu}(\overline{\phi_1} \vee \overline{\phi_2}),\ \neg(\mathtt{rev}\,(\overline{\nu})(\overline{\phi_1}) \vee \mathtt{rev}\,(\overline{\nu})(\overline{\phi_2}))) = (\overline{\nu}(\overline{\phi_1 \vee \phi_2}),\ \neg\mathtt{rev}\,(\overline{\nu})(\overline{\phi_1 \vee \phi_2})) =$$
$$(\overline{\nu}(\overline{\psi}),\ \neg\mathtt{rev}\,(\overline{\nu})(\overline{\psi})).$$

- The case in which $\psi = \phi_1 \wedge \phi_2$ is analogous to the latter case. $\square$