Modeling Paraconsistent Reasoning by Classical Logic

Ofer Arieli[1] and Marc Denecker[2]

[1] Department of Computer Science, The Academic College of Tel-Aviv
Antokolski 4, Tel-Aviv 61161, Israel
oarieli@mta.ac.il
[2] Department of Computer Science, University of Leuven
Celestijnenlaan 200A, B-3001 Heverlee, Belgium
marcd@cs.kuleuven.ac.be

Abstract. We introduce a general method for paraconsistent reasoning in knowledge systems by classical second-order formulae. 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 manage to overcome the shortcoming of classical logic in properly handling inconsistent data, and provide new ways of implementing multiple-valued paraconsistent reasoning in knowledge systems. 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).

1 Introduction

Any knowledge-based system for common-sense reasoning must be able to process incomplete and inconsistent information in a "proper" way. This implies, in particular, that (first-order) classical logic is inappropriate for such systems. 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 [7], therefore everything classically follows from a contradictory set of premises. If follows, therefore, that knowledge-based systems should use other (or more general) formalisms for handling uncertainty.

Preferential reasoning [23] 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 [5, 6] 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., [1, 2, 6, 20, 21]); 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 [1, 2], 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 [1, 2, 5, 6] 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, 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., [1–3, 11, 12, 20, 21]).
- 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 (implementations of) proof systems are known.

Our goal in this paper is to show a way in which these problems can be avoided (or at least alleviated) altogether. 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. Moreover, preference criteria on four-valued theories are translated into 'circumscriptive-like" formulae [17, 18], and thus paraconsistent reasoning may be automatically computed by some specialized methods for compiling circumscriptive theories (such as those described in [10, 22]), and incorporated into algorithms such as SCAN [19] and DLS [8, 9], for reducing second-order formulae to their first-order equivalents. [1]

2 Preliminaries

2.1 Preferential reasoning

First we briefly review the basic notions of preferential reasoning [23].

Definition 2.1. 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 $M \times \Sigma$ (called the *satisfaction relation*), and $\leq$ (the *preference relation*) is a binary relation on the elements of M.

Most often, the preference relation is a partial order or at least a pre-order (i.e., reflexive and transitive). In this paper this will always be the case.

Definition 2.2. Let $\mathcal{M} = (M, \models, \leq)$ be a preferential model, Γ a set of formulae in a language Σ, and $m \in M$. Then m *satisfies* Γ (notation: $m \models \Gamma$) if $m \models \gamma$ for every $\gamma \in \Gamma$. m *preferentially satisfies* Γ (alternatively, m is a $\leq$*-most preferred model* of Γ) if m satisfies Γ, and for each other $n \in M$ s.t. $n \leq m$ and n satisfies Γ, it holds that $m \leq n$. 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 2.3. 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 if every element of $!(\Gamma, \leq)$ satisfies ψ. We denote this by $\Gamma \models_{\leq} \psi$.

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 [17]. Furthermore, this approach is the semantical basis of some well-known general patterns for non-monotonic reasoning, introduced in [13–16], and it is a key concept behind many formalisms for nonmonotonic and paraconsistent reasoning (see, e.g., [1–3, 11, 12, 20, 21]). Our purpose in this paper is to propose techniques of expressing preferential reasoning by formulae in the underlying language. Next we define the framework for doing so.

[1] For a longer version of this paper see [4].

2.2 The underlying semantical structure

The formalism that we consider here is based on Belnap's four-valued algebraic structure [5, 6], denoted by $\mathcal{FOUR}$ (Figure 1). 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 a contradictory knowledge, and $\bot$ corresponds to an incomplete knowledge. This interpretation of the meaning of the truth values will be useful in what follows for modeling paraconsistent reasoning.[2] 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* that each element exhibits.

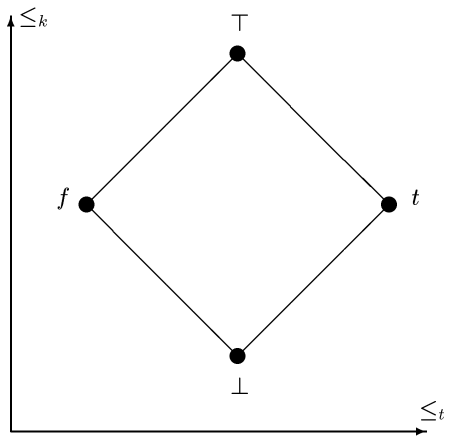

Fig. 1. $\mathcal{FOUR}$

In what follows we shall denote by $\wedge$ and $\vee$ the meet and join operations on $(FOUR, \leq_t)$. 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).

[2] This was also the original motivation of Belnap when he introduced $\mathcal{FOUR}$.

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 standard way:
$\nu(\psi \wedge \phi) = glb_{\leq_t}(\nu(\psi), \nu(\phi))$, $\nu(\psi \vee \phi) = lub_{\leq_t}(\nu(\psi), \nu(\phi))$, $\nu(\neg\psi) = \neg\nu(\psi)$.
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 $mod(\Gamma)$.

2.3 Four-valued preferential reasoning

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

Definition 2.4. 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 [24], i.e., it is reflexive, monotonic, and transitive. It is also shown there that $\models^4$ is paraconsistent, compact, and has cut-free, sound and complete Hilbert-type and Gentzen-type proof systems. However, 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 what follows we therefore refine the reasoning process by using the techniques discussed above for preferential reasoning. Below are some useful preference criteria.

Definition 2.5. [2] 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 the preference orders given in Definition 2.5 has its own rationality: according to $\leq_k$, for instance, one prefers valuations that reflect as minimal information as reasonably possible. 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. Similarly, $\leq_{\{\top,\bot\}}$, prefers those valuations that are as classical as possible. I.e., those ones that assign classical truth values whenever possible.

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

Example 2.1. Let $\Gamma = \{p, \neg p \vee q, \neg q, r \vee q\}$. The four-valued models of Γ are given in Table 1.

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

Model	p	q	r	Model	p	q	r
M_1	$\top$	f	t	M_6	t	$\top$	$\top$
M_2	$\top$	f	$\top$	M_7	$\top$	$\top$	$\bot$
M_3	t	$\top$	$\bot$	M_8	$\top$	$\top$	f
M_4	t	$\top$	f	M_9	$\top$	$\top$	t
M_5	t	$\top$	t	M_{10}	$\top$	$\top$	$\top$

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 2.5 induces a corresponding preferential consequence relation:

Definition 2.6. [2] 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.
- $\Gamma \models^4_{\{\top,\bot\}} \psi$ if every most classical model of Γ is a model of ψ.

Example 2.2. Consider again the set Γ of Example 2.1, 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 of Definition 2.6 are particular cases of the preferential entailment relations $\models_{\leq}$, given in Definition 2.3. It is also easy to see that all of these consequence relations are paraconsistent and have the following properties (see [2, 3] for further details):

1. $\Gamma \models^4_k \psi$ *iff* $\Gamma \models^4 \psi$.
   Thus $\models^4_k$ 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 $\models^4$.
2. *Denote by* $\models^2$ *the two-valued classical consequence relation. 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* $\Gamma \models^2 \psi$.

   Thus $\models^4_{\{\top,\bot\}}$ is equivalent to classical logic on consistent theories and is nontrivial w.r.t. inconsistent theories in the sense that not all formulas are entailed.

3 Paraconsistent classical reasoning

This section shows how to simulate paraconsistent reasoning by classical reasoning. We propose a transformation such that four-valued entailment for theories can be defined in terms of classical two-valued entailment for the transformed theories. Moreover, we show that four-valued preferential entailment can be defined in terms of classical entailment for the transformed theories augmented with circumscriptive axioms.

3.1 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: $t = (1,0)$, $f = (0,1)$, $\top = (1,1)$, $\bot = (0,0)$. In this representation the negation operator is defined in $\mathcal{FOUR}$ by $\neg(x,y) = (y,x)$, and the corresponding partial orders in $\mathcal{FOUR}$ are represented by the following rules: for every $x_1, x_2, y_1, y_2 \in \{0,1\}$,

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

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

It follows, in particular, that in the representation by pairs of two-valued components, the $\leq_t$-meet (i.e., the greatest lower bound w.r.t. $\leq_t$) and the $\leq_t$-join (i.e., the least upper bound w.r.t. $\leq_t$) in $\mathcal{FOUR}$ are defined as follows:

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

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

It is obvious that there is a one-to-one correspondence between four-valued valuations and pairs of two-valued valuations. We shall denote these pairs of two-valued components by $\nu = (\nu_1, \nu_2)$. So if, for instance, $\nu(\psi) = t$, then $\nu_1(\psi) = 1$ and $\nu_2(\psi) = 0$.

The preference criteria considered in the previous section may now be reformulated as follows:

Lemma 3.1. *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* $\nu_2(p) \leq \mu_2(p)$.
- $\nu \leq_{\{\top\}} \mu$ *iff for every atom* p, *if* $\nu_1(p) \wedge \nu_2(p) = 1$, *then* $\mu_1(p) \wedge \mu_2(p) = 1$ *as well.*

– $\nu \leq_{\{\top,\bot\}} \mu$ *iff for every atom* p, *if* $(\nu_1(p) \wedge \nu_2(p)) \vee (\neg\nu_1(p) \wedge \neg\nu_2(p)) = 1$, *then* $(\mu_1(p) \wedge \mu_2(p)) \vee (\neg\mu_1(p) \wedge \neg\mu_2(p)) = 1$ *as well.*

Proof: Immediately follows from the corresponding definitions. □

Given the language Σ, define the language $\Sigma^{\pm}$ based on the alphabet consisting of symbols p^+, p^- for each atom p of Σ. Let ψ be a formula in Σ. Denote by $\widehat{\psi}$ 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),[3] then substituting every occurrence in ψ' of an atomic formula p that is not preceded by a negation, by the new predicate symbol p^+, and replacing every other occurrence of p in ψ', together with the negation that precedes it, by the new predicate symbol p^-. For instance, if $\psi = \neg(p \vee \neg q)$, then $\psi' = \neg p \wedge q$, and so $\widehat{\psi} = p^- \wedge q^+$. Given a theory Γ, we shall write $\widehat{\Gamma}$ for the set $\{\widehat{\psi} \mid \psi \in \Gamma\}$. Note that $\widehat{\Gamma}$ can be obtained from Γ in a linear time. Moreover, for every Γ, $\widehat{\Gamma}$ is a positive theory, and hence it is consistent.

Given a four-valued valuation $\nu = (\nu_1, \nu_2)$, $\widehat{\nu}$ denotes the two-valued valuation on $\Sigma^{\pm}$, defined by $\widehat{\nu}(p^+) = \nu_1(p)$ and $\widehat{\nu}(p^-) = \nu_2(p)$. Extensions to complex formulae in $\Sigma^{\pm}$ are defined in the usual way. [4]

Definition 3.1. Given a valuation $\nu = (\nu_1, \nu_2)$, denote $\overline{\nu} = (\neg\nu_2, \neg\nu_1)$.

Proposition 3.1. $\nu(\psi) = (\widehat{\nu}(\widehat{\psi}),\ \neg\widehat{\overline{\nu}}(\widehat{\psi}))$.

Proof: Let ψ' be the negation normal form of ψ. Since ψ and ψ' are logically equivalent in $\mathcal{FOUR}$, $\nu(\psi)$ is the same as $\nu(\psi')$. The rest of the proof is by an induction on the structure of ψ':

$\psi' = p$: $(\widehat{\nu}(\widehat{p}), \neg\widehat{\overline{\nu}}(\widehat{p})) = (\widehat{(\nu_1, \nu_2)}(p^+), \neg\widehat{(\neg\nu_2, \neg\nu_1)}(p^+)) = (\nu_1(p), \neg\neg\nu_2(p)) = (\nu_1(p), \nu_2(p)) = \nu(p)$.

$\psi' = \neg p$: $(\widehat{\nu}(\widehat{\neg p}), \neg\widehat{\overline{\nu}}(\widehat{\neg p})) = (\widehat{(\nu_1, \nu_2)}(p^-), \neg\widehat{(\neg\nu_2, \neg\nu_1)}(p^-)) = (\nu_2(p), \neg\neg\nu_1(p)) = (\nu_2(p), \nu_1(p)) = \neg(\nu_1(p), \nu_2(p)) = \neg\nu(p) = \nu(\neg p)$.

$\psi' = \phi_1 \vee \phi_2$: $\nu(\phi_1 \vee \phi_2) = \nu(\phi_1) \vee \nu(\phi_2) = (\widehat{\nu}(\widehat{\phi_1}), \neg\widehat{\overline{\nu}}(\widehat{\phi_1})) \vee (\widehat{\nu}(\widehat{\phi_2}), \neg\widehat{\overline{\nu}}(\widehat{\phi_2})) = (\widehat{\nu}(\widehat{\phi_1}) \vee \widehat{\nu}(\widehat{\phi_2}), \neg\widehat{\overline{\nu}}(\widehat{\phi_1}) \wedge \neg\widehat{\overline{\nu}}(\widehat{\phi_2})) = (\widehat{\nu}(\widehat{\phi_1} \vee \widehat{\phi_2}), \neg(\widehat{\overline{\nu}}(\widehat{\phi_1}) \vee \widehat{\overline{\nu}}(\widehat{\phi_2}))) = (\widehat{\nu}(\widehat{\phi_1} \vee \widehat{\phi_2}), \neg\widehat{\overline{\nu}}(\widehat{\phi_1} \vee \widehat{\phi_2})) = (\widehat{\nu}(\widehat{\phi_1 \vee \phi_2}), \neg\widehat{\overline{\nu}}(\widehat{\phi_1 \vee \phi_2}))$.

The case $\psi' = \phi_1 \wedge \phi_2$ is similar to that of $\phi_1 \vee \phi_2$. □

---

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

[4] 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 $\nu'(p) = (\nu(p^+), \nu(p^-))$.

3.2 Simulating (preferential) four-valued reasoning by classical logic

In what follows we use the pairwise representations, considered in the previous section, for the following goals:

1. Showing that four-valued reasoning can be simulated by classical reasoning,
2. Constructing circumscriptive formulae for defining four-valued preferential reasoning.

For item (1) above we first need the following lemma:

Lemma 3.2. *For every four-valued valuation ν and a formula ψ in Σ, $\nu(\psi)$ is designated iff $\widehat{\nu}(\widehat{\psi})=1$.*

Proof: $\nu(\psi)$ is designated iff $\nu_1(\psi)=1$, iff (Proposition 3.1) $\widehat{\nu}(\widehat{\psi})=1$. □

The following result is an immediate corollary of Lemma 3.2:

Theorem 3.1. *$\Gamma \models^4 \psi$ iff $\widehat{\Gamma} \models^2 \widehat{\psi}$.*

It follows, therefore, that four-valued reasoning may be implemented by two-valued theorem provers. Moreover, since $\widehat{\Gamma}$ 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 the classical entailment.

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

Corollary 3.1. *In the language without negations, $\Gamma \models^4 \psi$ iff $\Gamma \models^2 \psi$.*

We turn now to the preferential case. To extend the above technique to deal with preferential four-valued reasoning, we must express that the encoded four-valued interpretation is minimal with respect to a preference relation $\leq$. This can be accomplished by introducing a circumscription axiom. The proviso is that we are able to express the preference relation $\leq$ objectively, by a formula $\Psi_{\leq}$. The first point to check out is therefore how to express a semantical preference relation $\leq$ in an axiom.

Let $\boldsymbol{p} = \{p_1, .., p_n\}$ be the set symbols of our language Σ, and let $\boldsymbol{p}^{\pm}$ be the set of symbols $\{p_1^+, p_1^-, \ldots, p_n^+, p_n^-\}$. 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 $\boldsymbol{q}$ as renaming of the symbols of $\boldsymbol{p}$. With $(\boldsymbol{p}^{\pm}\!:\!\nu; \boldsymbol{q}^{\pm}\!:\!\mu)$, we denote the two-valued interpretation that interprets symbols p_i^+ as $\nu_1(p)$, p_i^- as $\nu_2(p)$, q_i^+ as $\mu_1(p)$ and q_i^- as $\mu_2(p)$ for every $1 \leq i \leq n$.

Definition 3.2. A preferential order $\leq$ is *represented* by a formula $\Psi_{\leq}(\boldsymbol{p}^{\pm}, \boldsymbol{q}^{\pm})$ if for every four-valued valuations ν and μ we have that $\nu \leq \mu$ iff $(\boldsymbol{p}^{\pm}\!:\!\nu, \boldsymbol{q}^{\pm}\!:\!\mu)$ satisfies $\Psi_{\leq}(\boldsymbol{p}^{\pm}, \boldsymbol{q}^{\pm})$.

Given a formula ψ of Σ, denote by $\widehat{\psi}(\boldsymbol{P}^{\pm})$ the formula that is obtained by substituting symbols $\boldsymbol{P}^{\pm}$ for symbols $\boldsymbol{p}^{\pm}$ in $\widehat{\psi}$.

Proposition 3.2. *Let $\Psi_{\leq}(\boldsymbol{p}^{\pm}, \boldsymbol{P}^{\pm})$ 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 $\widehat{\nu}$ satisfies $\widehat{\psi}$ and the following formula:*

$$\mathsf{Circ}_{\leq}(\boldsymbol{p}^{\pm}) \;=\; \forall(\boldsymbol{P}^{\pm})\,\{\,\widehat{\psi}(\boldsymbol{P}^{\pm}) \;\rightarrow\; (\Psi_{\leq}(\boldsymbol{P}^{\pm}, \boldsymbol{p}^{\pm}) \rightarrow \Psi_{\leq}(\boldsymbol{p}^{\pm}, \boldsymbol{P}^{\pm}))\,\}.$$

Proof: By Corollary 3.2, ν is a model of ψ iff $\widehat{\nu}$ satisfies $\widehat{\psi}$. It remains to show that the fact that $\widehat{\nu}$ 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, $\widehat{\nu}$ 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 3.1. Denote the formula $\bigwedge_{i=1}^{n} ((p_i^+ = P_i^+) \wedge (p_i^- = P_i^-))$ by $\boldsymbol{p}^{\pm} = \boldsymbol{P}^{\pm}$,[5] and let $\Psi_{<}(\boldsymbol{P}^{\pm}, \boldsymbol{p}^{\pm})$ be the following formula: $\Psi_{\leq}(\boldsymbol{P}^{\pm}, \boldsymbol{p}^{\pm}) \wedge \neg\Psi_{\leq}(\boldsymbol{p}^{\pm}, \boldsymbol{P}^{\pm})$. Then:

a) The formula $\mathsf{Circ}_{\leq}(\boldsymbol{p}^{\pm})$ of Proposition 3.2 may be rewritten as follows:

$$\forall(\boldsymbol{P}^{\pm})\ \{\ \widehat{\psi}(\boldsymbol{P}^{\pm}) \;\rightarrow\; \neg\Psi_{<}(\boldsymbol{P}^{\pm}, \boldsymbol{p}^{\pm})\ \}$$

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

$$\forall(\boldsymbol{P}^{\pm})\,\{\,[\,\widehat{\psi}(\boldsymbol{P}^{\pm}) \wedge \Psi_{\leq}(\boldsymbol{P}^{\pm}, \boldsymbol{p}^{\pm})\,] \rightarrow \boldsymbol{p}^{\pm} = \boldsymbol{P}^{\pm}\,\}$$

The following theorem is an immediate corollary of Proposition 3.2:

Theorem 3.2. *Let Γ be a set of formulae and ψ a formula in Σ. Let $\mathsf{Circ}_{\leq}$ be the formula given in Proposition 3.2 for a preferential relation $\leq$. Then $\Gamma \models_{\leq} \psi$ iff $\widehat{\Gamma} \cup \mathsf{Circ}_{\leq} \models^2 \widehat{\psi}$.*

Proposition 3.2 gives a general characterization in terms of "formula circumscription" [18] 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 3.2. Again, this can be done in a polynomial time.

Next we define formulae that represent the preferential relations considered above.

Definition 3.3. In what follows we shall write $x \preceq y$ for $x \rightarrow y$, and $x \prec y$ for $(x \rightarrow y) \wedge \neg(y \rightarrow x)$.

[5] In the context of two-valued logic, $p = q$ denotes $\forall \overline{x}.p(\overline{x}) \leftrightarrow q(\overline{x})$.

Lemma 3.3. *Let n be the number of different atomic formulae in Σ. Then:*

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

$$\Psi_{\le_k}(\boldsymbol{p}^{\pm}, \boldsymbol{P}^{\pm}) \;=\; \bigwedge_{i=1}^{n} ((p_i^+ \preceq P_i^+) \wedge (p_i^- \preceq P_i^-))$$

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

$$\Psi_{\le_{\{\top\}}}(\boldsymbol{p}^{\pm}, \boldsymbol{P}^{\pm}) \;=\; \bigwedge_{i=1}^{n} ((p_i^+ \wedge p_i^-) \;\preceq\; (P_i^+ \wedge P_i^-))$$

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

$$\Psi_{\le_{\{\top\}}}(\boldsymbol{p}^{\pm}, \boldsymbol{P}^{\pm}) = \bigwedge_{i=1}^{n} (((p_i^+ \wedge p_i^-) \vee (\neg p_i^+ \wedge \neg p_i^-)) \;\preceq\; ((P_i^+ \wedge P_i^-) \vee (\neg P_i^+ \wedge \neg P_i^-)))$$

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

$$\begin{aligned}
\nu \le_k \mu &\iff \forall 1 \le i \le n \;\; \nu(p_i) \le_k \mu(p_i) \\
&\iff \forall 1 \le i \le n \;\; \nu_1(p_i) \le \mu_1(p_i) \text{ and } \nu_2(p_i) \le \mu_2(p_i) \\
&\iff (\boldsymbol{p}^{\pm}:\nu,\, \boldsymbol{P}^{\pm}:\mu) \text{ satisfies } \textstyle\bigwedge_{i=1}^{n} ((p_i^+ \preceq P_i^+) \wedge (p_i^- \preceq P_i^-)) \\
&\iff (\boldsymbol{p}^{\pm}:\nu,\, \boldsymbol{P}^{\pm}:\mu) \text{ satisfies } \Psi_{\le_k}(\boldsymbol{p}^{\pm}, \boldsymbol{P}^{\pm}).
\end{aligned}$$

□

By Proposition 3.2, Lemma 3.3(a), and Note 3.1(b), we have the following corollary:

Corollary 3.2. *A valuation $\nu = (\nu_1, \nu_2)$ is a k-minimal model of ψ iff $\widehat{\nu}$ satisfies $\widehat{\psi}$ and $\mathsf{Circ}_{\le_k}(\boldsymbol{p}^{\pm})$, where $\mathsf{Circ}_{\le_k}(\boldsymbol{p}^{\pm})$ is the following formula:* [6]

$$\forall(\boldsymbol{P}^{\pm}) \; \{ \; [\widehat{\psi}(\boldsymbol{P}^{\pm}) \wedge \bigwedge_{i=1}^{n} ((P_i^+ \preceq p_i^+) \wedge (P_i^- \preceq p_i^-))] \rightarrow [\bigwedge_{i=1}^{n} ((P_i^+ = p_i^+) \wedge (P_i^- = p_i^-))] \; \}$$

As in Corollary 3.2, the most consistent models and the most classical models of a given theory can be represented by formulae of the form $\mathsf{Circ}_{\le_{\{\top\}}}(\boldsymbol{p}^{\pm})$ and $\mathsf{Circ}_{\le_{\{\top,\bot\}}}(\boldsymbol{p}^{\pm})$, obtained by respectively integrating the formulae given in parts (b) and (c) of Lemma 3.3 with $\mathsf{Circ}_{\le}$, given in Proposition 3.2.

In the remaining of this section we consider a uniform way of representing $\mathsf{Circ}_{\le_{\{\top\}}}(\boldsymbol{p}^{\pm})$, $\mathsf{Circ}_{\le_{\{\top,\bot\}}}(\boldsymbol{p}^{\pm})$, and some other formulae that correspond to preferential criteria like $\le_{\{\top\}}$ and $\le_{\{\top,\bot\}}$. For this, let $\Delta \subseteq FOUR$. Define an order relation $<_\Delta$ on $FOUR$ by $x <_\Delta y$ iff $x \notin \Delta$ while $y \in \Delta$. A corresponding

[6] Note that $\mathsf{Circ}_{\le_k}(\boldsymbol{p}^{\pm})$ is a standard circumscriptive axiom in the sense of [17].

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:

Definition 3.4. For $\Delta \subseteq FOUR$, let $\Lambda_\Delta(p^+, p^-) = \bigvee_{x \in \Delta} \Lambda_x(p^+, p^-)$, where

$$\Lambda_t(p^+, p^-) = p^+ \wedge \neg p^-, \qquad \Lambda_f(p^+, p^-) = \neg p^+ \wedge p^-,$$

$$\Lambda_\bot(p^+, p^-) = \neg p^+ \wedge \neg p^-, \qquad \Lambda_\top(p^+, p^-) = p^+ \wedge p^-. \text{ [7]}$$

Similar arguments as those in Lemma 3.3 show that the formula

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

represents the preferential relation $\leq_\Delta$. Therefore, by Proposition 3.2,

Proposition 3.3. *A valuation ν is a $\leq_\Delta$-preferred model of ψ iff $\widehat{\nu}$ satisfies $\widehat{\psi}$ and the following formula:*

$$\mathsf{Circ}_{\leq_\Delta}(\boldsymbol{p}^\pm) \;=\; \forall(\boldsymbol{P}^\pm)\,\{\, \widehat{\psi}(\boldsymbol{P}^\pm) \;\rightarrow\; (\Psi_{\leq_\Delta}(\boldsymbol{P}^\pm, \boldsymbol{p}^\pm) \rightarrow \Psi_{\leq_\Delta}(\boldsymbol{p}^\pm, \boldsymbol{P}^\pm))\,\}.$$

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 3.2, 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 [8, 9], available at `http://www.ida.liu.se/labs/kplab/projects/dls/circ.html`). [8]

- 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 $\widehat{\Gamma} = \{Q^+(a), Q^+(b), Q^-(a)\}$. Circumscribing $\widehat{\Gamma}$ where Q^+ and Q^- are simultaneously minimized, yields the following result:

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

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

[8] In what follows we deliberately consider very simple cases. Our experience is that for more complex theories the output quickly becomes more complicated, and so not comprehensible by humans (it *is* manageable in automated computations, though).

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 information may be important from analysis point of view.

– Suppose that in the previous example one wants to impose the law of excluded middle. 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 $\widehat{\psi} = \forall x(Q^+(x) \vee Q^-(x))$. Circumscribing $\widehat{\Gamma} \cup \{\widehat{\psi}\}$ 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 $\widehat{\psi}$ 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 $\Gamma \cup \{\psi\}$.

5 Conclusion

In this paper we have introduced a method for paraconsistent reasoning in knowledge systems by classical second-order formulae. Our method touches upon several important aspects. First, it shows that two-valued reasoning may be useful for simulating inference procedures in the context of many-valued semantics. Second, this approach demonstrates the usefulness of circumscription not only as a general method for non-monotonic reasoning, but also as an appealing technique for implementing paraconsistent reasoning. Finally, this is another evidence to 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). This enables a potentially wide area for practical implementations. For instance, we have shown that preferential multiple-valued reasoning can be incorporated with practical applications for automated reasoning and theorem proving.

References

1. O.Arieli, A.Avron. *The logical role of the four-valued bilattice.* Proc. LICS'98, pp.218–226, IEEE Press, 1998.
2. O.Arieli, A.Avron. *The value of the four values.* Artificial Intelligence 102(1), pp.97–141, 1998.

3. 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.
4. O.Arieli, M.Denecker. *Circumscriptive approaches to paraconsistent reasoning.* Technical Report CW-304, Department of Computer Science, University of Leuven, 2001.
5. N.D.Belnap. *A useful four-valued logic.* Modern Uses of Multiple-Valued Logic (G.Epstein, J.M.Dunn, editors), pp.7–37, Reidel Publishing Company, 1977.
6. N.D.Belnap. *How computer should think.* Contemporary Aspects of Philosophy (G.Ryle, editor), pp.30–56, Oriel Press, 1977.
7. N.C.A.da-Costa. *On the theory of inconsistent formal systems.* Notre Dame Journal of Formal Logic 15, pp.497–510, 1974.
8. P.Doherty, W.Lukaszewicz, A.Szalas. *Computing circumscription revisited: Preliminary report.* Proc. IJCAI'95, pp.1502–1508, 1995.
9. P.Doherty, W.Lukaszewicz, A.Szalas. *Computing circumscription revisited: A reduction algorithm.* Journal of Automated Reasoning 18, pp.297–334, 1997.
10. M.L.Ginsberg. *A circumscriptive theorem prover.* Artificial Intelligence 29, pp.209–230, 1989.
11. M.Kifer, E.L.Lozinskii. *RI: A logic for reasoning with inconsistency.* Proc. LICS'89, pp.253–262, IEEE Press, 1989.
12. M.Kifer, E.L.Lozinskii. *A logic for reasoning with inconsistency.* Journal of Automated Reasoning 9(2), pp.179–215, 1992.
13. S.Kraus, D.Lehmann, M.Magidor. *Nonmonotonic reasoning, preferential models and cumulative logics.* Artificial Intelligence 44(1–2) pp.167–207, 1990.
14. D.Lehmann, M.Magidor. *What does a conditional knowledge base entail?* Artificial Intelligence 55, pp.1–60, 1992.
15. D.Makinson. *General theory of cumulative inference.* Non-Monotonic Reasoning (M.Reinfrank, editor), LNAI 346, pp.1–18, Springer, 1989.
16. 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.
17. J.McCarthy. *Circumscription – A form of non monotonic reasoning.* Artificial Intelligence 13(1–2), pp.27–39, 1980.
18. J. McCarthy. *Applications of circumscription to formalizing common-Sense knowledge.* Artifical Intelligence 28, pp.89–116, 1986.
19. H.J.Ohlbach, *SCAN – Elimination of predicate quantifiers.* Proc. CADE'96, (M.A.McRobbie, J.Slaney, editors), LNAI 1104, pp.161–165, Springer, 1996.
20. G.Priest. *Reasoning about truth.* Artificial Intelligence 39, pp.231–244, 1989.
21. G.Priest. *Minimally Inconsistent LP.* Studia Logica 50, pp.321–331, 1991.
22. T.Przymusinski. *An algorithm to compute circumscription.* Artificial Intelligence 38, pp.49–73, 1991.
23. Y.Shoham. *Reasoning about change.* MIT Press, 1988.
24. A.Tarski. *Introduction to logic.* Oxford University Press, N.Y., 1941.