Some semantic properties of Negative Rationality

Michael Freund

Departement de Mathématiques,
Université d'Orléans
45067 Orléans, France
freund@centre.univ-orleans.fr

Daniel Lehmann

Institute of Computer Science,
Hebrew University
Jerusalem 91904, Israel
lehmann@cs.huji.ac.il

Abstract

We study inference relations that satisfy the property of Negation Rationality. This rule expresses the fact that when a conclusion may be infered from a premisse, it cannot be withdrawn both when adding a new hypothesis to this premisse and when adding its negation. For preferential relations, this rule is known to be weaker than the one of Disjunctive Rationality, and therefore weaker than that of Rational Monotonicity. We provide a semantic characterization of preferential inference relations that satisfy the property of Negation Rationality, and show that, in the finite case, the set of preferential inference relations that satisfy Negation Rationality and admit an injective model is exactly the set of preferential inference relations that satisfy the property of Disjunctive Rationality.

Introduction

Among the non-Horn rules that were presented and studied in KLM and some following related papers, the rule of Negation Rationality can be considered as a basic one, in the sense that it is commonly agreed that it should be supported by any rational reasoner. This rule reflects indeed the fact that it is unreasonable to expect that a conclusion γ, legitimely drawn from a premisse α, should be withdrawn both when adding an additional information β and when adding the information $\neg\beta$. Thus, this property can be formalized on the form

(N.R.) If $\alpha \mid\sim \gamma$, **then, for any formula** β, **either** $\alpha\wedge\beta \mid\sim \gamma$, **or** $\alpha\wedge\neg\beta \mid\sim \gamma$.

which can be equivalently expressed by saying that **if** γ **cannot be induced by the** $\alpha\wedge\beta$ **and** γ **cannot be induced from** $\alpha\wedge\neg\beta$, **then** γ **cannot be induced from** α. To take an example proposed in KLM, if one does not expect a party to be great knowing that Peter is coming, and if one does not expect the party to be great knowing that Peter is not coming, then one cannot reasonably expect the

party to be great, since either Peter will attend the party or he will not.

There has been very little work done on negation rationality. In fact, most of the papers on preferential logics that dealt with non-Horn properties focused on **rationality**, that is on preferential relations that satisfy the property of **Rational Monotonicity**. The reason is first that it has been claimed that the only "reasonable" consequence relations are the rational ones. The arguments to this thesis may be found in KLM and LM. Besides, the models attached to rational relations are easy to handle and generate an attractive "kappa-calculus" that cannot be done in the general framework of preferential relations, even in the disjunctive case. This explains why, in the problem of default ranking or belief changes for instance, the solutions that are mostly looked for are of a rational type.

Nevertheless the study of inference relations satisfying Negation Rationality is interesting ***per se*** and completes the work previously done on preferential relations. As will be shown, preferential relations that satisfy this condition can be represented by a certain class of preferential models, and a soundness and completness result for this family of relations can be established, which is analogue to the ones that were obtained in the case of preferential, disjunctive or rational relations. Moreover, we will prove that, in the case of a logically finite language, the preferential relations that satisfy Negation Rationality and that admit an injective model are exactly the ones that satisfy the condition of Disjunctive Rationality. This result is to be compared with similar ones obtained in FLM for relations satisfyting the condition of **Weak Contraposition**. We do not know yet if it can be extended in the general case where nothing is supposed on the underlying language.

We suppose the reader familiar with the basic results of preferential logics, which we will only recall briefly to make this paper self-contained. A full presentation and justification of these results can be found in KLM, LM or F.

Background

We let $\mathfrak{L}$ be the set of well formed formulas over a set of propositional variables, using the classical propositional connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$ and $\leftrightarrow$. With the language $\mathfrak{L}$, semantics is provided by the set $\mathfrak{C}$ of all assignments of truth values to the propositional variables. Elements of $\mathfrak{C}$ will be refered to as **worlds** and the satisfaction relation between a world m and a formula α is defined as usual and written $m \models \alpha$.
Thus, $m \models \alpha \vee \beta$ iff $m \models \alpha$ or $m \models \beta$, and $m \models \neg\alpha$ iff **not**$(m \models \alpha)$.

For every subset A of $\mathfrak{L}$, we write $m \models A$ iff m satisfies all the elements of A.

The classical consequence operation attached to $\mathfrak{L}$ and $\mathfrak{C}$ will be denoted by Cn: for any subset A of $\mathfrak{L}$, Cn(A) is then the set of all formulas α of $\mathfrak{L}$ such that $m \models \alpha$ for all worlds m that

satisfy A. If $\alpha \in \mathfrak{L}$ we write Cn(α) for Cn({α}).

Recall that Cn is a Tarski consequence operation: it satisfies

1) Reflexivity $A \subseteq Cn(A)$
2) Idempotence $Cn[Cn(A)] = Cn(A)$
3) Monotonicity If $A \subseteq B$, then $Cn(A) \subseteq Cn(B)$.

Moreover, Cn is **compact**: For any subset A of $\mathfrak{L}$ and $\alpha \in Cn(A)$, there exists a **finite** subset A_0 of A such that $\alpha \in Cn(A_0)$.

A subset A of $\mathfrak{L}$ is called **consistent** if $Cn(A) \neq \mathfrak{L}$ or, equivalently if there exists a world m such that m satisfies A. The compactness of Cn implies that A is consistent if and only if all finite subsets of A are consistent.

Preferential inference relations

Following K.L.M., we call **preferential inference relation** on $\mathfrak{L}$ a relation $|\sim$ that satisfies the six following rules:

1) Reflexivity
$\alpha \mid\sim \alpha$
2) Left Logical Equivalence
If $Cn(\alpha) = Cn(\beta)$ and $\alpha \mid\sim \gamma$, then $\beta \mid\sim \gamma$
3) Right Weakening
If $\beta \in Cn(\alpha)$ and $\gamma \mid\sim \alpha$, then $\gamma \mid\sim \beta$
4) Cut
If $\alpha\wedge\beta \mid\sim \gamma$ and $\alpha \mid\sim \beta$, then $\alpha \mid\sim \gamma$
5) Or
If $\alpha \mid\sim \gamma$ and $\beta \mid\sim \gamma$, then $\alpha\vee\beta \mid\sim \gamma$
6) Cautious Monotonicity
If $\alpha \mid\sim \beta$ and $\alpha \mid\sim \gamma$, then $\alpha\wedge\beta \mid\sim \gamma$.

These properties are shared by all preferential relations. Some other rules of a different type (non-Horn clauses) were defined and studied in KLM, LM and FLM. We recall them briefly:

Rational Monotonicity
(R.M.) If $\alpha \mid\sim \beta$ and **not** ($\alpha \mid\sim \neg\gamma$), then $\alpha\wedge\gamma \mid\sim \beta$.

Disjunctive rationality
(**D.R**) If $\alpha\vee\beta \mid\sim \gamma$ and **not** ($\alpha \mid\sim \gamma$), then $\beta \mid\sim \gamma$.

Weak Contraposition
(**W.C.**) If $\gamma\wedge\alpha \mid\sim \beta$ and **not** ($\gamma \mid\sim \beta$), then $\gamma\wedge\neg\beta \mid\sim \neg\alpha$.

Given such an inference relation, we shall denote by $C_{\sim}(\alpha)$ – or $C(\alpha)$ when there is no ambiguity– the set of all consequences of a formula α, that is the set of all β's such that $\alpha \mid\sim \beta$. Rule **3**, if satisfied, implies then that $C(\alpha)$ is closed w.r.t. Cn, that is
$Cn[C(\alpha)] = C(\alpha)$.

Preferential models

In KLM the authors defined a *preferential model* in the following way:

Let **W** = (**S**,$\prec$,l) be a triple where $\prec$ is an irreflexive and transitive relation defined on a set **S** and l is a mapping from **S** in the set of worlds $\mathfrak{C}$.

For any element α of $\mathfrak{L}$, denote by α^* be the set of all elements s of **S** such that s "satisfies" α, i.e. such that $l(s) \models \alpha$. Then, **W** is said to be a **preferential model** if the following condition of **smoothness** is satisfied: for any $\alpha \in \mathfrak{L}$ and s in α^*, either s is minimal in α^* or there exists an element t in **S**, $t \prec s$, such that t is minimal in α^* .

Such a preferential model defines a preferential relation $|\sim_W$ by:

(def) $\alpha \,|\sim_W \beta$ iff all minimal elements of α^* satisfy β.

As in [], let us write $s \models_\prec \alpha$ iff s satisfies α and s is minimal in α^*. Then the preferential relation induced by the preferential model W can be written

$\alpha \,|\sim_W \beta$ iff $s \models_\prec \alpha$ implies $s \models \beta$.

It was proved in KLM that, conversely, for any preferential relation $|\sim$ there exists a preferential model **W** such that $|\sim = |\sim_W$. The model built by the authors for this purpose, and which we will use later, is called the **canonical model of $|\sim$.** It is defined as follows: the set of states **S** is the set of pairs (m,α) where m is a world, and α a formula such that $m \models \alpha$; the label function l associates to every state (m,α) the world m. As for the relation "$\prec$", it is a strict partial order on **S** defined by $(m,\alpha) \prec (n,\beta)$ iff m does not satisfy β and $\alpha \vee \beta \,|\sim \alpha$, this last condition being equivalent to the condition that $C(\alpha \vee \beta) = C(\alpha)$.

The condition of Negation Rationality

An inference relation $|\sim$ is said to satisfy Negation Rationality if the following property holds for every formulas α, β and γ:

(**N.R.**) **If** $\alpha \,|\sim \gamma$, **then either** $\alpha \wedge \beta \,|\sim \gamma$, **or** $\alpha \wedge \neg\beta \,|\sim \gamma$.

Note that if the given relation is supposed to satisfy Cautious Monotonicity, the property (**N.R.**) is always satisfied for formulas α such that $C(\alpha)$ is a maximal consistent set, since in this case one has either $\alpha \,|\sim \beta$ or $\alpha \,|\sim \neg\beta$.

Using the notation introduced above, the condition of Negation Rationality is equivalent to $C(\alpha) \subseteq C(\alpha \wedge \gamma) \cup C(\alpha \wedge \neg\gamma)$. It is possible in most cases to simplify it. We can indeed use the following result:

Lemma 1 ***Let X,Y and Z be theories such that $X \subseteq Y \cup Z$. Then $X \subseteq Y$ or $X \subseteq Z$.***

Proof: Suppose X is not a subset of Y and let x be an element of X that is not an element of Y. Then $x \in Z$. If x' is any other element of X, one has $x \wedge x' \in X$. But $x \wedge x' \notin Y$ since $x \notin Y$ and Y is closed. Therefore $x \wedge x' \in Z$, so $x' \in Z$ and we have shown that $X \subseteq Z$. ■

Suppose now that the relation $|\sim$ satisfies **Right Weakening**, that is that $C(\phi)$ is a theory for all formulas ϕ. Then, using the above lemma, one sees that Negation Rationality is equivalent to the condition

(7) $C(\alpha) \subseteq C(\alpha \wedge \beta)$ or $C(\alpha) \subseteq C(\alpha \wedge \neg\beta)$.

If now we suppose moreover, that C is a **preferential** inference relation, one has the property of **Conditionalization**: for all formulas ϕ and ψ, $C(\phi \wedge \psi) \subseteq Cn[\phi, C(\psi)]$, and condition (**7**) is therefore equivalent to

(8) $C(\alpha \wedge \beta) = Cn[\beta, C(\alpha)]$ or $C(\alpha \wedge \neg\beta) = Cn[\neg\beta, C(\alpha)]$

Although the condition of Negation Rationality is a weak non-Horn condition – it is known to be strictly weaker than Disjunctive Rationality, which is itself strictly weaker than Rational Monotonicity, it is not in general satisfied in most of the known nonmonotonic logics. It was shown for instance in KLM that circumscriptive reasoning does not necessarily obey Negation Rationality. In fact the same (negative) result holds for Reiter default logic and Poole systems:

Observation 1 ***There exists a set D of default rules such that the associated sceptical inference relation C does not satisfy Negation Rationality.***

Proof: We take an example taken from Makinson []. Let **D** consist of the two default rules $d = (\textbf{true}; \beta/\beta)$ and $d' = (\alpha \vee \beta; \neg\beta/\neg\beta)$ where α and β are two formulas such that $\alpha \vee \beta$ is not a tautology and α is consistent with $\neg\beta$. Note that the only Reiter extension of **true** under **D** is $E = Cn(\beta)$, so one has **true** $|\sim \beta$. On the other hand, $\{\alpha \vee \beta\}$ has two extensions, E and $E' = Cn\{\alpha, \neg\beta\}$. Since $\beta \notin E \cap E'$, one does not have $\alpha \vee \beta \; |\sim \beta$. But one does not have either $\neg(\alpha \vee \beta) \; |\sim \beta$, therefore, setting $\gamma = \alpha \vee \beta$, one sees that **true** $|\sim \beta$ while neither $\gamma \; |\sim \beta$ nor $\neg\gamma \; |\sim \beta$.■

Observation 2 (Makinson, personal communication)***There exists a set D of defaults such that the associated Poole inference relation does not satisfy Negation Rationality.***

Proof: We recall that the Poole inference relation associated with a set **D** of defaults is given by
$C(\alpha) = Cn(\alpha, D_\alpha)$, where D_α is the intersection of all the maximal subsets of **D** that are consistent with α.
For the example that we are looking for, we take for set of defaults $\textbf{D} = \{(\alpha \vee \beta) \wedge \gamma; \neg\alpha \wedge \gamma; (\neg\alpha \vee \beta) \wedge \neg\gamma; \alpha \wedge \neg\gamma\}$. It is the easy to check that $\beta \in C(\textbf{true})$, $\beta \notin C(\alpha)$ and $\beta \notin C(\neg\alpha)$.■

It is well known that the inference relations generated "simple" Poole systems, (i.e. systems like the preceding one with finite set of defaults and without constraint), are preferential. It has been proved moreover that these relations can be represented by **injective** preferential models, that is models with an injective label function, where the set of states may be taken as a subset of the set of worlds. The above observation shows

then that **even in the case where the language is supposed to be logically finite, there exist preferential relations that admit an injective model but do not satisfy Negation Rationality.** This point will be developed later.

Comparison with some other non-Horn rules.

In KLM, LM and FLM, it was proved that, in the preferential case, Negation Rationality is strictly weaker than Disjunctive Rationality but is incomparable with Weak Contraposition, and that Weak Contraposition and Negation Rationality together do not imply Disjunctive Rationality, even in the case of a finite language. It may be interesting to ask what kind of link exists, if any, between Negation Rationality and Weak Rationality, since this latter rule is part of the "basic postulates" of AGM. Recall first that Weak Rationality is the rule:

(**W.K.**) **If true** $|\sim \alpha$, **then either true** $|\sim \neg\beta$, **or** $\beta |\sim \alpha$.

It is not difficult to show that Negation Rationality does not imply Weak Rationality, as can be seen from the following example taken from L.M.,lemma 3.3: the undelying language has three propositional variables p, q and r; the preferential relation $|\sim$ is defined by a model with four states, $s_1 \prec s_2$ and $s_3 \prec s_4$, where the two top states satisfy the propositional variables p and q, the state s_1 satisfies p and r and the state s_3 satisfies q and r. This relation $|\sim$ was shown to satisfy Negation Rationality. We claim it does not satisfy Weak Rationality. Indeed we have **true** $|\sim \neg p \vee \neg q$, but we do not have **true** $|\sim \neg p$ (since s_1 is minimal and satisfies p) and we do not have either $p |\sim \neg p \vee \neg q$ (since s_3 is minimal in p^* and satisfies $p \wedge q$).

The following example shows that Weak Rationality does not imply Negation Rationality. Let p, q, r, and s be propositional variables, and s_0, s_1, s_2, s_3, and s_4 five states such that s_0 satisfies only s (in {p, q, r, s}), s_1 satisfies p, q and r, s_2 sa&tisfies only r, s_3 satisfies q and r and s_4 satisfies p and r. The order is given by s_0 is a minimum, $s_1 \prec s_2$ and $s_3 \prec s_4$. One sees then easily that $r |\sim q$ and that neither $r \wedge p |\sim q$, nor $r \wedge \neg p |\sim q$, so the relation $|\sim$ does not satisfy Negation Rationality. Let us prove that it satisfies Weak Rationality: Suppose that α and β are formulas such that $\alpha \in C(\mathbf{true})$ and $\neg\beta \notin C(\mathbf{true})$. We have to show that $\beta |\sim \alpha$. Let t be a minimal state in β^*. We prove that $t \models \alpha$. Since $\neg\beta \notin C\mathbf{true})$ we have that s_0 satisfies β. But s_0 is a minimum, so $s_0 = t$ or $s_0 \prec t$. The latter cannot hold by the choice of t. Therefore t = s and t satisfies α as desired.

Putting these results together, we obtain the

Observation 3 ***Negation Rationality is incomparable to Weak Rationality.***

Some weak non-Horn clauses reveal to be equivalent to stronger ones when the underlying language is supposed to be logically **finite.** This is the case for Weak Contraposition, that turns out then to be equivalent to Rational Monotonicity (FLM, th.3.10). The corresponding result for Negation Rationality is the following

Theorem 1 ***Let C be a preferential relation defined on a finite***

language. Suppose that C admits an injective model. Then if C satisfies Negation Rationality, it satisfies Disjunctive Rationality.

Proof: Suppose that C does not satisfy Disjunctive Rationality. Then there exist formulas α, β and γ such that $\gamma \in C(\alpha \vee \beta)$, $\gamma \notin C(\alpha)$ and $\gamma \notin C(\beta)$.

Since $\gamma \notin C(\alpha)$, there exists a world n′ minimal in α^* such that $n' \models \neg\gamma$. This world is therefore not minimal in $\alpha \vee \beta$, and there exists a world $n \prec n'$ that is minimal in $(\alpha \vee \beta)^*$. Note that n is then minimal in β^* and that $n \models \gamma$.

Similarly, there exist two worlds m′ and m such that $m \prec m'$, m′ minimalin β^*, m minimal in α^*, $m' \models \neg\gamma$ and $m \models \gamma$.

We claim that $m' \models \alpha$. Otherwise, m′ would be minimal in $(\beta \wedge \neg\alpha)^*$ and since m′ does not satisfy γ, we would have $\gamma \notin C(\beta \wedge \neg\alpha)$ as well as $\gamma \notin C(\alpha)$, contradicting the fact that $\gamma \in C(\alpha \vee \beta)$ and that C satisfies Negation Rationality. Similarly, we have $n' \succeq \beta$.

Let now δ' be a formula such that $n' \succeq \delta'$ and $m' \succeq \neg\delta'$. Such a formula always exists since n′ and m′ are different worlds (n′ is minimal in α^* but m′ is not). Set $\delta = \alpha \wedge \beta \wedge '$. Note that n′ is the only world in {m, n, m′, n′} that satisfies δ. Let λ be the formula that is true only in one of the worlds m, n, m′, n′. Set $\mu = (\alpha \wedge \neg\beta) \vee (\beta \wedge \neg\alpha)$ and $\eta = (\neg\beta \vee \delta) \wedge (\beta \vee \neg\delta)$. Then we have that

1) $\mu \in C(\lambda)$.

Indeed m and n are the only worlds minimal in λ^*, and they satisfy μ.

2) $\mu \notin C(\lambda \wedge \eta)$.

For n′ is minimal in $\lambda \wedge \eta$ but n′ does not satisfy μ.

3) $\mu \notin C(\lambda \wedge \neg\eta)$.

For m′ is minimal in $\lambda \wedge \neg\eta$ and does not satisfy μ.

It follows from this that C does not satisfy Negation Rationality, and the proof of the theorem is therefore completed.■

As an easy consequence of the theorem, we can sharpen the relations between Negation Rationality and Weak Contraposition:

Corollary ***There exist preferential relations defined on a finite language that can be represented by an injective model and that satisfy Negation Rationality but that do not satisfy Weak Contraposition.***

Proof: It follows from **theorem 1** that the set of all preferential relations defined on a finite language satisfying Negation Rationality and admitting an injective model is identical to the set of prefrential relations defined on that language that satisfy Disjunctive Rationality. On the other hand, the set of preferential relations defined on a finite language that admit an injective model and that satisfy Weak Contraposition is known to be the set of all rational relations on that language. The corollary follows from the fact that preferential relations that satisfy Disjunctive Rationality can always be represented by an injective model (F) and that the condition of Rationality is strictly stronger than that of Disjunctive Rationality, even in the finite case.■

Semantics

When defining a subfamily of preferential relations, an interesting problem is that of the possibility of representing it by suitable preferential models. In KLM, the authors introduced a class of "canonical" models that were suitable to represent all the preferential relations. It happens that the models built to represent preferential relations that satisfy disjunctive rationality or rational monotonicity were not taken as a subclass of the canonical models. It turns out, though, that the subfamily of preferential relations satisfying negation rationality can be represented by canonical models. More precisely, we have the following

Observation 4 ***Let C be a preferential relation that satisfies negation rationality and W = {S,l,≺} its canonical model. Let*** α ***and*** β ***be two formulas. Then one at least of the two following properties holds: either for any state s minimal in*** $(\alpha\wedge\beta)^*$ ***there exists a state s′ minimal in*** α^* ***such that l(s) = l(s′), or for any state t minimal in*** $(\alpha\wedge\neg\beta)^*$***, there exists a state t′ minimal in*** α^* ***such that l(t) = l(t′).***

Proof: Suppose that C satisfies negation rationality and let α and β be two formulas. By (**1**), one has $C(\alpha) \subseteq C(\alpha\wedge\beta)$ or $C(\alpha) \subseteq C(\alpha\wedge\neg\beta)$. Suppose we are in the first case, and let $s = (n,\phi)$ be a state minimal in $(\alpha\wedge\beta)^*$. Then $n = l(s)$ satisfies $C(\alpha)$. Let $s' = (n,\alpha)$.
We have $s' \in$ **S**, $l(s') = l(s)$ and s' is minimal in α^* as desired. The proof in the case $C(\alpha) \subseteq C(\alpha\wedge\neg\beta)$ is similar.■

We now prove the converse of the above result: suppose that C is a preferential relation that is defined by a model **W**, not necessarily canonical, that enjoys the property stated in **theorem 1**. We show that C satisfies negation rationality. Let α and β be two formulas, and $\gamma \in C(\alpha)$. Suppose that $\gamma \notin C(\alpha\wedge\beta)$. We have to prove that $\gamma \in C(\alpha\wedge\neg\beta)$. By hypothesis, there exists a state s minimal in $(\alpha\wedge\beta)^*$ such that l(s) does not satisfy γ. Note that there cannot exist a state s′ minimal in α^* with $l(s) = l(s')$. since otherwise l(s′) would satisfy $C(\alpha)$, hence γ. The hypothesis made on **W** implies then that for any state t minimal in $(\alpha\wedge\neg\beta)^*$ there exists a state t′ minimal in α^* such that $l(t) = l(t')$. We have $l(t') \models \gamma$, and therefore $l(t) \models \gamma$. This being true for any state t minimal in $(\alpha\wedge\neg\beta)^*$ shows that $\gamma \in C(\alpha\wedge\neg\beta)$ as claimed.
This result together with **Observation 4** gives a representation theorem for preferential relations satisfying negation rationality:

Theorem 2 ***A preferential relation satisfies negation rationality if and only if it can be represented by a model where, for any formulas*** α ***and*** β***, either for each state s minimal in*** $(\alpha\wedge\beta)^*$ ***there exists a state s′ minimal in*** α^* ***such that l(s) = l(s′), or for each state t minimal in*** $(\alpha\wedge\neg\beta)^*$ ***there exists a state t′ minimal in*** α^* ***such that l(t) = l(t′).***

Conclusion and further work