Computer Science Department Technical Report
University of California
Los Angeles, CA 90024-1596

ON THE RELATION BETWEEN RATIONAL CLOSURE AND SYSTEM-Z

Moises Goldszmidt
Judea Pearl

July 1991
CSD-910043

In proceedings of Third International Workshop on Nonmonotonic Reasoning.
May 31 - June 3, 1990, pp. 130-140.

TECHNICAL REPORT
R-139
MAY 1990

On the Relation Between Rational Closure and System-Z*

Moisés Goldszmidt
moises@cs.ucla.edu

Judea Pearl
judea@cs.ucla.edu

Cognitive Systems Lab.
Dept. of Computer Science, UCLA
Los Angeles, CA 90024

Abstract

Recent progress towards unifying the probabilistic and model-preference semantics for nonmonotonic reasoning has led to two systems of ranking. The first, called *system-Z* ([Pearl 90]), ranks sentences according to a *consistency* based criterion. The second called *rational closure* ([Lehmann 89]), ranks models according to a *preference*-based criterion. In this paper we show that the entailment relation defined by these two systems are the same. Additionally, we provide a procedure for deciding entailment that requires a polynomial number of propositional satisfiability tests, and discuss the adequacy of this entailment relation in defeasible reasoning applications.

1 Background: System-Z and Rational Closure

1.1 Preliminary definitions

Let L be a closed set of well formed propositional formulas, built in the usual way from a *finite* set of propositional variables and the connectives "$\vee$" and "$\neg$". Small case greek letters α, β and γ will be used to denote formulas in L.

A world w is an assignment of truth values to the propositional variables in L. Hence, if there are n propositional variables in L there will be 2^n worlds. Let $\mathcal{U}$ stand for the set

*This work was supported in part by National Science Foundation grant #IRI-88-21444 and Naval Research Laboratory grant #N00014-89-J-2007.

of worlds. The satisfaction of a formula α by a world ω is defined as usual, and will be written as $w \models \alpha$.

Using the binary connective "$\rightarrow$" and two formulas α and β from L we can construct the conditional sentence $\alpha \rightarrow \beta$. We will use Δ to denote a set of conditional sentences. A conditional sentence $\alpha \rightarrow \beta$ will be *verified* by w, if $w \models \alpha$ and $w \models \beta$. The same sentence will be *falsified* or *violated* by w, if $w \models \alpha$ but $w \not\models \beta$. If $w \not\models \alpha$, the sentence will be considered as neither verified nor falsified. We define the relation of *tolerance* based on these notions:

Definition 1 (Tolerance.) *Let Δ be a set of conditional assertions $\alpha_i \rightarrow \beta_i$ for $0 < i < n$. Δ is said to* tolerate *the sentence $d : \alpha_n \rightarrow \beta_n$ written $T(d||\Delta)$, if and only if $\exists w \in \mathcal{U}$ such that $w \models \alpha_n \wedge \beta_n$ and $w \models \bigwedge_i \alpha_i \supset \beta_i$* [1] *for all i such that $0 < i < n$.*

Thus, a conditional sentence d is *tolerated* by a set Δ if we can find a world w that verifies d while no other sentence in Δ is falsified by w. The formula $\alpha \supset \beta$ will be called the *material counterpart* of $\alpha \rightarrow \beta$. Note that if $\alpha \rightarrow \beta$ is not falsified by w, then its material counterpart is satisfied by w. We will say that a non-empty set Δ of conditional sentences is *confirmable* if we can find a sentence $d \in \Delta$ that is tolerated by Δ. Consistency and 0-entailment are defined below:

Definition 2 (Consistency.) *A set Δ is* consistent *if and only if every nonempty subset Δ' of Δ is confirmable.*

Thus, Δ is consistent if and only if we can find a tolerated sentence in every subset Δ' of Δ.

Definition 3 (0-entailment.) *Given a consistent set Δ and a conditional sentence d : $\alpha \rightarrow \beta$. We will say that Δ* 0-entails *d, written $\Delta \models_0 d$, if and only if $\Delta \cup \{\alpha \rightarrow \neg\beta\}$ is inconsistent.*

These definitions of consistency and entailment can be proven as theorems if the sentences in Δ are interpreted as conditional probabilities arbitrarily close to 1. Consistency assures the existence of a satisfying probabilistic model for Δ. The 0-entailment relation guarantees that conclusions will receive arbitrarily high probability values in all probabilistic models of Δ in which the premises also receive arbitrarily high probability values. For an early motivation of probabilistic interpretation for conditional sentences the reader is referred to [Adams 75] and more recently [Pearl 88] and [Geffner 89], in the context of defeasible reasoning. Extensions of consistency and entailment to sets containing both defeasible and strict sentences can be found in [Goldszmidt & Pearl 90]. The closure defined by 0-entailment contains the maximal set of "safe" conclusions that can be drawn from

[1] The symbol "$\supset$" denotes material implication.

Δ, namely, conclusions that remain undefeasible under augmentations of Δ by additional sentences, as long as the database remains consistent. This closure was was proposed in [Pearl 89a] as a *conservative core* that ought to be common to all nonmonotonic formalisms, and proven equivalent to the preferential closure of Δ, where the sentences in Δ are given a *preferential* interpretation ([Lehmann & Magidor 88]). Due to its extremely conservative nature, 0-entailment does not properly handle *irrelevant* features. For example, if Δ consists solely of the sentence $a \to b$, we are not able to conclude $a \wedge c \to b$, where a, b and c are propositional variables in L. We now introduce two formalisms, system-Z and rational closure, which extend the inferential power of 0-entailment.

1.2 System Z

The condition of confirmability required by the definition of consistency leads to a natural ordering of the sentences in Δ. Given a set Δ, we first identify every sentence that is tolerated by Δ, assign to each such sentence the label 0 and remove them from Δ. Next, we attach the label 1 to every sentence that is tolerated by the remaining ones and so on. Continuing in this manner, we build an ordered partition of $\Delta = (\Delta_0, \Delta_1, \ldots, \Delta_K)$, where

$$\Delta_i = \{\alpha \to \beta \mid T(\alpha \to \beta || \Delta - \Delta_0 - \ldots - \Delta_{i-1})\} \tag{1}$$

The label attached to each sentence in the partition defines the Z-ordering. The process of constructing this partition also amounts to testing the consistency of Δ since it terminates with a complete partition if and only if Δ is consistent. The number of propositional satisfiability tests required to perform the partition is bounded by $|\Delta|^2$, and consequently this procedure is polynomial for sublanguages in which propositional satisfiability is also polynomial (e.g. Horn clauses see [Dowling & Gallier 84]).

The only case in which this process will fail to complete the partition, is if it reaches a nonconfirmable subset Δ_u, i.e. a subset in which no sentence is tolerated. We assign an ∞ label to these sentences and denote this subset by Δ_∞. We introduce the definition of an inconsistent formula:

Definition 4 (Inconsistent formulas.) *A formula α is said to be* inconsistent *with respect to a set Δ (or* Δ-inconsistent*), iff $\alpha \to True$ is not tolerated by any subset in the partition of Δ.*

Note that any unconfirmable set must be a subset of Δ_∞ and it can be shown from Def. 4 that the antecedents of sentences in Δ_∞ are all Δ-inconsistent formulas.

Based on this ordering we can now define three ranking functions: on the sentences in Δ, on the worlds in $\mathcal{U}$, and on arbitrary formulas in L [2]. Given a sentence $\alpha \to \beta \in \Delta$, its *rank* will be equal to i, if and only if $\alpha \to \beta \in \Delta_i$ (if Δ is inconsistent, i might be ∞). The

[2] For the sake of simplicity we will use the term *rank* to denote all three functions.

rank of a world $w \in \mathcal{U}$, will be the smallest integer n such that all sentences having *rank* higher or equal to n are not falsified by w. Finally, the *ranking* of a formula $\alpha \in L$ as the lowest *rank* of all worlds satisfying α, or $rank(\alpha) = \infty$ if α is Δ-inconsistent. The total order imposed by these ranking functions can be interpreted as *preferences* among worlds (or states of affairs). It is in fact the (unique) lowest ranked preference ordering on worlds that satisfies Δ, if every sentence in Δ is interpreted as a partial order preferring worlds verifying $\alpha \to \beta$ over those falsifying it ([Pearl 90]). Similarly, the notion of entailment proposed below will proclaim $\alpha \to \beta$ a plausible (or rational) consequence of Δ, if the ranking function induced by Δ prefers worlds verifying $\alpha \to \beta$ over those falsifying it.

Definition 5 (1-entailment.) *A conditional sentence $\alpha \to \beta$ is said to be* 1-entailed *by Δ, written $\Delta \models_1 \alpha \to \beta$, if and only if*

$$rank(\alpha \wedge \beta) < rank(\alpha \wedge \neg\beta) \tag{2}$$

or if $rank(\alpha) = \infty$.

Thus, a conditional sentence $\alpha \to \beta$ is 1-entailed by Δ iff there exists an integer i such that the set of sentences with *rank* higher than i tolerates $\alpha \to \beta$ but does not tolerate $\alpha \to \neg\beta$. Note that, once the partition of Δ is known, verifying 1-entailment takes $O(\log K)$ satisfiability tests, where K is the number of partition sets in the Z-ordering.

Further ramifications of system Z are explored in [Pearl 90].

1.3 The rational closure

We now review the definitions of *ranked models* and *rational closure*. The reader is referred to [Lehmann & Magidor 88] and [Lehmann 89] for details and motivation.

Definition 6 (Ranked Models.) *A* ranked model *W is a triple $\langle S, l, \prec\rangle$ where S is a set of states, l is a function mapping each $s \in S$ to a world $\omega \in \mathcal{U}$, and for which the ordering relation $\prec$ may be defined in the following way: there is a totally ordered set Ω (the strict order of Ω will be denoted by $<$) and a function $r : S \mapsto \Omega$ such that $s \prec t$ if and only if $r(s) < t(s)$. A state S is said to* satisfy *a formula α iff $l(s) \models \alpha$.*

Additionally, $\prec$ is required to satisfy the following smoothness condition: $\forall \alpha \in L$, the set of states $\hat{\alpha} \stackrel{\text{def}}{=} \{s | s \in S, l(s) \models \alpha\}$ is smooth. Where if $V \subseteq U$; V is *smooth* if and only if $\forall t \in V$, either $\exists s$ minimal in V, such that $s \prec t$ or t is itself minimal in V.

A ranked model W will constitute a model for a set Δ of conditional sentences if and only if for every $\alpha \to \beta \in \Delta$, it is true that for any state s minimal in $\hat{\alpha}$, $l(s) \models \beta$.

Let Δ^p denote the set of conditionals $\alpha \to \beta$ such that $\alpha \to \beta$ is satisfied by all ranked models of Δ; Δ^p constitutes the preferential closure of Δ. It is proven in [Lehmann & Magidor 88] that, given a consistent Δ, $\alpha \rightarrow \beta \in \Delta^p$ if and only if $\Delta \models_0 \alpha \rightarrow \beta$. Thus, Δ^p is equivalent to 0-entailment.

In [Lehmann 89], a more powerful closure of Δ, called *rational closure* (denoted here by Δ^r), is introduced. The proposal involves closing Δ^p under a rule called *rational monotony*: If $\alpha \rightarrow \beta$ is in Δ^r and $\alpha \rightarrow \neg\gamma$ is not, then $\alpha \wedge \gamma \rightarrow \beta$ must be in Δ^r. This rule of inference allows the strengthening of the antecedent α of a conditional sentence by γ, whenever γ is not "atypical" relative to that sentence.

A more procedural description of the rational closure rests on the following definition which was first introduced in [Lehmann & Magidor 88] in the completeness proof between ranked models and ranked consequence relations:

Definition 7 (Exceptional Formulas.) *Let Δ be a set of conditional sentences, and let α, β be formulas from L that are not Δ-inconsistent. We will say that α is* more exceptional *than β in the context of Δ if and only if*

$$\alpha \vee \beta \rightarrow \neg\alpha \in \Delta^p \tag{3}$$

We can now assign a *degree* to each formula $\alpha \in L$ according to the following inductive procedure ([Lehmann 89]): The state zero of the induction is the set of formulas of degree strictly less than zero is empty. Suppose that $i \geq 0$ and that the set of formulas of *degree* less than i has been defined. The set of formulas of *degree* i are those formulas α that are not of *degree* less than i and that satisfy: $\forall\beta \in L$ such that α is more exceptional than β, β is of *degree* less than i. If this procedure fails to assign a *degree* to α, then $degree(\alpha) = \infty$.

Let Δ^r denote the rational closure of Δ, then $\alpha \rightarrow \beta \in \Delta^r$ if and only if $degree(\alpha) < degree(\alpha \wedge \neg\beta)$ or $degree(\alpha) = \infty$. Thus a conditional sentence $\alpha \rightarrow \beta$ is in the rational closure of Δ if and only if the state that satisfies both α and β is less exceptional than the state that satisfies α but does not satisfy β (or if α cannot be assigned a degree of exceptionality).

2 Equivalence of Rational Closure and 1-entailment

We will prove the equivalence between rational closure and 1-entailment for finite sets Δ, by first showing that $\forall\alpha \in L$, $degree(\alpha) = rank(\alpha)$. The $degree(\alpha)$ is defined in terms of Δ^p, while $rank(\alpha)$ is based on the notion of 0-entailment. These two forms of entailment were proven equivalent for the case of consistent Δ in [Lehmann & Magidor 88], and we now remove this restriction from the definition of 0-entailment in a way that is compatible with the probabilistic semantics of Δ [3]. This is accomplished by defining the conditional probability $P(\beta|\alpha) = 1$ if $P(\alpha) = 0$ (as done in [Adams 66] and more recently

[3]See previous section

in [Goldszmidt, Morris & Pearl 90]). Thus, given a finite set Δ we define an extension of 0-entailment called *probabilistic entailment* which does not require Δ to be consistent:

Definition 8 (Probabilistic entailment.) *Let Δ be a finite set of conditional sentences, Δ* probabilistically entails *$\phi \rightarrow \gamma$ iff for all $\varepsilon > 0$ there exists a $\delta > 0$ such that for all probability assignments P for L, if $P(\beta|\alpha) \geq 1-\varepsilon$ for all $\alpha \rightarrow \beta \in \Delta$ then $P(\gamma|\phi) \geq 1-\delta$.*

Thus, a sentence will be probabilistically entailed if and only if it can be assigned arbitrarily high probability values in all probability models in which the premises receive probability values arbitrarily close to one. This notion of entailment is similar to the one advocated in [Goldszmidt & Pearl 90], with the exception that P is not required to be positive for the antecedent of every conditional sentence (i.e. P is not constrained to be *proper*). As a consequence of forcing *improper* probabilities to be 1, Def. 8 will allow sentences of the form $\alpha \rightarrow False$ to be probabilistically entailed even when α is logically consistent.

We now show the equivalence of probabilistic entailment and preferential entailment[4], following the strategy used in [Lehmann & Magidor 88], where this result was proven in the case of consistent Δ.

Lemma 1 ([Lehmann & Magidor 88].) *Any conditional sentence preferentially entailed by Δ is probabilistically entailed by Δ.*

This lemma can be proven by showing that the rules for preferential entailment (see [Lehmann & Magidor 88]) are sound with respect to Def. 8 in the same way that soundness is proven in [Adams 75] or [Adams 66].

The next lemma asserts the converse of Lemma 1:

Lemma 2 *If Δ does not preferentially entail $\alpha \rightarrow \beta$, then Δ does not probabilistically entail $\alpha \rightarrow \beta$.*

Proof: (sketch) We use the following model construction[5]: given a ranked model W with a finite set of states S, we define a probability assignment P such that all states of the same rank will receive equal probability, and such that the weight w_n of a set of states of rank n, w_n will satisfy $\frac{w_{n+1}}{w_n} = \varepsilon$. There is exactly one probability assignment satisfying these requirements for each finite ranked model, and we can define a probability assignments on formulas in the natural way. It is clear that $P(\alpha) = 0$ iff the formula α is inconsistent in W, (i.e. W satisfies $\alpha \rightarrow False$). Let us consider the conditional probability $P(\beta|\alpha)$: if α is inconsistent then $P(\beta|\alpha) = 1$ by definition. Else if W satisfies $\alpha \rightarrow \beta$, $P(\beta|\alpha) > 1 - \varepsilon - \varepsilon^2 - \ldots$ will approach one as ε approaches zero. If, on the other hand, W does not satisfy $\alpha \rightarrow \beta$, the conditional probability cannot exceed $1 - \frac{1}{m}$ where m

[4] We say that a sentence is preferentially entailed, iff the sentence belongs to Δ^p.
[5] See [Adams 75], [Lehmann & Magidor 88].

is the number of states at the rank which is minimal for α, and the probability is bounded away from one as ε approaches zero. Given that $\alpha \rightarrow \beta$ is not preferentially entailed by Δ it is possible to find a ranked model W that satisfies Δ, satisfies $\alpha \rightarrow \neg\beta$ and for which α is not inconsistent [Lehmann 89]. We can now use the probability model construction outlined above and find a probability distribution in which all sentences in Δ are assigned probabilities arbitrarily close to one, $P(\alpha) > 0$, and since $\lim_{\varepsilon \to 0} P(\neg\beta|\alpha) = 1$, $P(\beta|\alpha)$ must approach zero and by Def. 8 $\alpha, \rightarrow \beta$ is not probabilistically entailed.

The following theorem provides an effective procedure for deciding whether a sentence is probabilistically entailed by Δ:

Theorem 1 *Let $\Delta' \stackrel{def}{=} \Delta \cup \{\alpha \rightarrow \neg\beta\}$. Then Δ probabilistically entails $\alpha \rightarrow \beta$ iff $\alpha \rightarrow \neg\beta \in \Delta'_\infty$*

Proof: Assume that $\alpha \rightarrow \neg\beta \notin \Delta'_\infty$. This implies that $\alpha \rightarrow \neg\beta$ belongs to some other partition set Δ'_k of Δ' and it is tolerated by the sentences in all rankings higher or equal to k. We can then build a probabilistic model in the manner outlined (in the discussion preceding Lemma 2), for which $P(\alpha) > 0$ and $\lim_{\varepsilon \to 0} P(\neg\beta|\alpha) = 1$ and consequently $\alpha \rightarrow \beta$ is not probabilistically entailed. To prove the *if* part of the theorem, we use the following result from [Adams 66]: a sentence $\alpha \rightarrow \beta$ is probabilistically entailed by a set Δ iff there exists a subset Δ_s such that any truth assignment falsifying $\alpha \rightarrow \beta$ must falsify at least one member of Δ_s, and any truth assignment not falsifying any member of Δ_s and verifying at least one member of Δ_s must verify $\alpha \rightarrow \beta$. Now if $\alpha \rightarrow \neg\beta \in \Delta'_\infty$ any truth assignment verifying $\alpha \rightarrow \neg\beta$ (i.e. falsifying $\alpha \rightarrow \beta$) must falsify at least one other sentence in Δ'_∞, or else, $\alpha \rightarrow \beta$ would obtain a finite rank. Moreover, any truth assignment not falsifying any member of Δ'_∞ and verifying at least one sentence in Δ'_∞ must falsify $\alpha \rightarrow \neg\beta$ (i.e. verify $\alpha \rightarrow \beta$), else this one sentence would obtain a finite rank. Thus, identifying Δ_s with Δ_∞, $\alpha \rightarrow \beta$ is probabilistically entailed by Δ.

Given that preferential entailment and probabilistic entailment are proven equivalent by Lemmas 1 and 2 we will refer to them indistinctively as p-entailment and use the symbol "$\models_p$". Theorem 1 asserts that a sentence $\alpha \rightarrow \beta$ is p-entailed by Δ, if and only if adding $\alpha \rightarrow \neg\beta$ to Δ, renders α Δ-inconsistent. The following is an easy corollary of Theorem 1:

Corollary 1 *Given a finite set Δ, p-entailment can be decided in polynomial time relative to propositional satisfiability.*

All lemmas and theorems below assume a finite set Δ and a Z-ranking function on Δ denoted $rank_\Delta$.

Lemma 3 *$rank_\Delta(\alpha) = 0$ iff $\forall\beta \in L$, $\alpha \vee \beta$ is consistent with respect to $\Delta' \stackrel{def}{=} \Delta \cup \{\alpha \vee \beta \rightarrow \alpha\}$*

Proof: If $rank_\Delta(\alpha) = 0$ then any word w_α of minimal ranking that satisfies α must satisfy the material counterpart of all sentences in Δ, and therefore must also verify $\alpha \vee \beta \rightarrow \alpha$,

$\forall \beta \in L$. Thus, $\alpha \vee \beta$ is not inconsistent with respect to Δ'. On the other hand, assume that $\forall \beta \in L$ $\alpha \vee \beta$ is consistent with respect to Δ', and $rank_\Delta(\alpha) > 0$. Let β' be a propositional tautology (i.e. equivalent to $True$); Δ' becomes $\Delta \cup \{True \rightarrow \alpha\}$, and it can only be confirmed by a world satisfying α (the sentence $True \rightarrow \alpha$ is falsified by any other world). But then either $rank(\alpha)$ must be equal to zero or Δ' is unconfirmable and therefore identical to Δ'_∞, a contradiction.

Lemma 4 *$\forall \alpha, \beta \in L$ such that α and β are not Δ-inconsistent. If $rank_\Delta(\beta) \geq rank_\Delta(\alpha)$ then $\alpha \vee \beta$ is consistent with respect to $\Delta' \stackrel{def}{=} \Delta \cup \{\alpha \vee \beta \rightarrow \alpha\}$*

Proof: Let α and β be two non Δ-inconsistent formulas, and let $rank_\Delta(\alpha) = i$ and $rank_\Delta(\beta) = i'$ with $i' \geq i$ (note that neither i nor i' can be ∞). For all $j < i$ the partition sets Δ_j and Δ'_j must be identical since every world w such that $rank_\Delta(w) < i$ does not satisfy either α or β and consequently $\alpha \vee \beta \rightarrow \beta$ is not falsified by any of these worlds. Now consider a world w', such that $rank_\Delta(w') = i$, and $w' \models \alpha$. This world verifies $\alpha \vee \beta \rightarrow \alpha$ and does not falsify any sentence with $rank_\Delta > i$. Thus, it follows that $\alpha \vee \beta \rightarrow \alpha \in \Delta_i$ and $i \neq \infty$.

Lemma 5 *$\forall \alpha \in L$, and $0 < m < \infty$. If $rank_\Delta(\alpha) = m$, then $\exists \beta \in L$ of $rank_\Delta(\beta) = m-1$, such that $\alpha \vee \beta$ is inconsistent with respect to $\Delta' = \Delta \cup \{\alpha \vee \beta \rightarrow \alpha\}$.*

Proof: Note that it is enough to show that there is an unconfirmable subset Δ'_I of Δ' such that $\alpha \vee \beta \rightarrow \alpha \in \Delta'_I$. Let β be the disjunction of all formulas $\gamma_i \in L$, such that $rank_\Delta(\gamma_i) = m - 1$, and each γ_i is the antecedent or consequence of some sentence in Δ. Let Δ_{m-1} be the subset of Δ containing only sentences of $rank_\Delta$ $m-1$. Clearly $\Delta_{m-1} \cup \{\alpha \vee \beta \rightarrow \alpha\}$ is not confirmable since any world verifying a sentence in Δ_{m-1} must falsify $\alpha \vee \beta \rightarrow \alpha$ and vice versa (note that otherwise $rank(\alpha) < m$).

Theorem 2 *$\forall \alpha \in L$ and $n \geq 0$; $rank(\alpha) = n$ if and only if $degree(\alpha) = n$*

Proof: If $rank(\alpha) = \infty$, α must be a Δ-inconsistent formula, $\Delta \models_p \alpha \vee \beta \rightarrow \neg\alpha$ $\forall \beta \in L$, and consequently the procedure described in the previous section will fail to assign a degree to α (i.e. $degree(\alpha) = \infty$). On the other hand if $degree(\alpha) = \infty$, then $\Delta \models_p \alpha \rightarrow False$ (see [Lehmann 89]) and it follows that α is Δ-inconsistent (i.e. $rank(\alpha) = \infty$). For $n \neq \infty$ the proof proceeds by induction on n. The basic case is essentially lemma 3. We assume that the theorem holds for $n < m$ and show that it is in fact true for $n = m$. If $rank(\alpha) = m$, then from lemma 4 it follows that $\forall \beta \in L$, if $\Delta \models_p \alpha \vee \beta \rightarrow \neg\alpha$ then $rank(\beta) < m$. Thus from the induction hypothesis, $degree(\beta) < m$ for all such β, and from the definition of degree of previous section $degree(\alpha) = m$. Similarly, if $degree(\alpha) = m$, it follows that $\forall \beta \in L$, if $\Delta \models_p \alpha \vee \beta \rightarrow \neg\alpha$ then $degree(\beta) < m$. By the induction hypothesis, $rank(\beta) < m$ and $rank(\alpha) \geq m$. Now, from lemma 5, if $rank(\alpha) > m$ then there must exist a β of $rank(\beta) = m$ such that $\Delta \models_p \alpha \vee \beta \rightarrow \neg\alpha$. Thus, $rank(\alpha)$ must be equal to m.

Corollary 2 *Given a finite set Δ, $\forall \alpha, \beta \in L$, $\alpha \rightarrow \beta \in \Delta^r$ if and only if $\Delta \models_1 \alpha \rightarrow \beta$.*

3 Discussion

The notion of 1-entailment (and consequently rational closure) has attractive computational features (assuming a finite set Δ). We present below a decision procedure for 1-entailment which is polynomial relative to propositional satisfiability:

Input: finite set Δ and a sentence $\alpha \rightarrow \beta$.

Output: answer YES/NO depending on whether $\Delta \models_1 \alpha \rightarrow \beta$.

1. TEST whether $\alpha \rightarrow \beta$ is tolerated by Δ.
2. TEST whether $\alpha \rightarrow \neg\beta$ is tolerated by Δ.
3. CASES indexed by results from TEST1-TEST2:
   (a) IF YES-YES or NO-YES then return NO.
   (b) IF YES-NO then return YES.
   (c) IF NO-NO then let $\Delta = \Delta_t \cup \Delta_u$, where Δ_t are the sentences tolerated by Δ and $\Delta_u = \Delta - \Delta_t$. If Δ_t is empty, then return YES; else call the procedure recursively with Δ_u replacing Δ.

The intuition behind the procedure is as follows: if both sentences are tolerated, or if $\alpha \rightarrow \beta$ is not tolerated but its negation ($\alpha \rightarrow \neg\beta$) is, then it follows that $rank(\alpha \wedge \beta) \geq rank(\alpha \wedge \neg\beta)$ and $\Delta \not\models_1 \alpha \rightarrow \beta$ (case 3.a). If, on the other hand $\alpha \rightarrow \beta$ is tolerated, but its negation is not, then $\Delta \models_1 \alpha \rightarrow \beta$ since $rank(\alpha \wedge \beta) < rank(\alpha \wedge \neg\beta)$ (case 3.b). If neither is tolerated and the set is unconfirmable, then the sentence is trivially 1-entailed since its antecedent is Δ-inconsistent; else we must perform the test again since neither sentence can be ranked at the current level and the procedure continues recursively. Note that, once we have the *rank* of all sentences in Δ, deciding 1-entailment for an arbitrary query $\alpha \rightarrow \beta$ requires at most $2 * (1 + \log K)$ satisfiability tests (using a binary search strategy), where K is the number of ranks in Δ. We simply determine the lowest *rank* k such that all sentences ranked k or higher tolerate $\alpha \rightarrow \beta$, repeat for $\alpha \rightarrow \neg\beta$, and compare.

The computational convenience of 1-entailment and rational closure is overshadowed by some counterintuitive inferences produced by these systems. Although rational closure properly handles irrelevant facts, i.e. if p is a proposition not appearing in Δ and $\alpha \rightarrow \beta \in \Delta^r$, then $\alpha \wedge p \rightarrow \beta \in \Delta^r$ (see [Lehmann 89]), it is incapable of sanctioning property inheritance across exceptional subclasses ([Pearl 90]). The problem is that the rank (or "abnormality") of a given world is completely determined by the sentence of highest rank falsified in that world. The next example illustrates this point: consider $\Delta = \{a \rightarrow$

$b, c \rightarrow d\}$, we can verify that $\Delta \not\models_1 (a \wedge \neg b \wedge c) \rightarrow d$ (because $rank(a \wedge \neg b \wedge c \wedge d) = rank(a \wedge \neg b \wedge c \wedge \neg d)$), yet we have no reason to believe that the abnormality affecting $a \rightarrow b$ is of any relevance to $c \rightarrow d$. Another problem with rational closure is the triggering or spurious conclusions by irrelevant elaborations. Consider $\Delta = \{a \rightarrow \neg b, c \rightarrow b\}$, naturally we cannot conclude $a \wedge c \rightarrow b$ or $a \wedge c \rightarrow \neg b$. Now we elaborate on the properties of c and add to Δ the information $\Delta' = \{c \rightarrow d, c \rightarrow \neg e, d \rightarrow e\}$; this elaboration suddenly renders c more exceptional than a, disturbs their symmetry, and sanctions the counterintuitive conclusion $\Delta \cup \Delta' \models_1 a \wedge c \rightarrow \neg b$. These problems are overcome by a formalism based on infinitesimal probabilities augmented by the principle of maximum entropy [Goldszmidt, Morris & Pearl 90], in which a more refined ordering is induced, sensitive to the number of rules tolerating a formula and not merely to their rank orders.

Counterexamples to the maximum entropy formalism, mostly dealing with the nature of causation, can also be constructed. The problem with both these systems is not the function used to rank the models, but rather their commitment to a *total* rank order among worlds. In fact, it seems that any proposal for defeasible reasoning that properly respects specificity and commits to a total order among worlds is bound to produce counterintuitive conclusions (by elaborations similar to those in the previous paragraph). The partial order proposed in [Geffner 89] seems to overcome these difficulties, at the expenses of an added complexity and a greater departure from its probabilistic origin.

4 Acknowledgements.

We are thankful to Tom Verma for useful discussions and imaginative counterexamples. Hector Geffner advised us against formalisms that commit to total orders.

References

[Adams 66] Adams, E., Probability and The Logic of Conditionals, in *Aspects of Inductive Logic*, ed. J. Hintikka and P. Suppes, Amsterdam: North Holland.

[Adams 75] Adams, E., *The Logic of Conditionals*, chapter II, Dordrecht, Netherlands: D. Reidel.

[Dowling & Gallier 84] Dowling, W. and J. Gallier, Linear-Time Algorithms for Testing the Satisfiability of Propositional Horn Formulae, *Journal of Logic Programming*, 3:267–284, 1984.

[Geffner 89] Geffner, H., Default Reasoning: Causal and Conditional Theories, UCLA Cognitive Systems Laboratory, PhD Dissertation, also Technical Report 137.

[Goldszmidt & Pearl 90] Goldszmidt, M. and J. Pearl, On The Consistency of Defeasible Databases, submitted to *Artificial Intelligence*.

[Goldszmidt, Morris & Pearl 90] Goldszmidt M., P. Morris and J. Pearl, The Maximum Entropy of Nonmonotonic Systems, to appear in AAAI-90 Proceedings, Boston 1990.

[Kraus et.al. 88] Kraus, S., D. Lehmann and M. Magidor, Preferential Models and Cumulative Logics, Technical Report TR 88-15, Dept. of Computer Science, Hebrew University, Jerusalem, Israel, November 1988.

[Lehmann 89] Lehmann, D., What Does a Knowledge Base Entail?, Proceedings of the First International Conference on Knowledge Representation, Toronto, Canada 1989, pp. 212–222.

[Lehmann & Magidor 88] Lehmann, D. and M. Magidor, Rational Logics and their Models: A Study in Cumulative Logics, TR-8816 Dept. of Computer Science, Hebrew Univ., Jerusalem, Israel.

[Pearl 88] Pearl, J., *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, chapter 10, Morgan Kaufmann Publishers Inc.

[Pearl 89a] Pearl, J., Probabilistic Semantics for Nonmonotonic Reasoning: A Survey, in Proceedings of the First Intl. Conf. on Principles of Knowledge Representation and Reasoning, Toronto, Canada, May 1989, pp. 505–516.

[Pearl 90] Pearl, J., System Z: A Natural Ordering of Defaults with Tractable Applications to Nonmonotonic Reasoning, in Proceedings TARK-90, M. Vardi (ed.), 1990, pp. 121–135.