Preferential Logics: the Predicate Calculus case *

(extended abstract)

Daniel Lehmann[†] and Menachem Magidor[‡]

Abstract

Suppose a knowledge base contains information on how the world generally behaves and in particular contains the information that *birds, normally fly*. Suppose that we obtain the information that Tweety is a bird, why should we conclude that it is plausible that Tweety flies? The answer to this question is unexpectedly sophisticated since the *obvious* substitution rule has to be rejected. Our answer to this question is based on an extension to predicate calculus of the ideas presented in [7]. Preferential consequence relations over predicate calculi are defined. In addition to the rules satisfied by those relations in the propositional case, they satisfy two rules dealing with quantifiers. These rules are not enough to enable us to conclude that Tweety flies. The rational closure construction defined in [7] should be generalized to the predicate calculus case and, in the rational closure, Tweety should fly.

1 Introduction

Many systems that exhibit nonmonotonic behavior have been described and studied already in the literature. The general notion of nonmonotonic reasoning, though, has almost always been described only negatively, by the property it does not enjoy, i.e., monotony. We study here general patterns

*This work was partially supported by the Jean and Helene Alfassa fund for research in Artificial Intelligence and by a grant from the Basic Research Foundation, Israel Academy of Sciences and Humanities

[†]Department of Computer Science, Hebrew University, Jerusalem 91904 (Israel)

[‡]Department of Mathematics, Hebrew University, Jerusalem 91904 (Israel)

of nonmonotonic reasoning and try to isolate properties that could help us map the field of nonmonotonic reasoning by reference to positive properties. We concentrate on nonmonotonic consequence relations, defined in the style of Gentzen [3]. Both proof-theoretic and semantic points of view are developed in parallel.

Nonmonotonic logic is the study of those ways of inferring additional information from given information that do not satisfy the monotony property satisfied by all methods based on classical (mathematical) logic. In Mathematics, if a conclusion is warranted on the basis of certain premises, no additional premises will ever invalidate the conclusion. In everyday life, however, it seems clear that we, human beings, draw sensible conclusions from what we know and that, on the face of new information, we often have to take back previous conclusions, even when the new information we gathered in no way made us want to take back our previous assumptions. For example, we may hold the assumption that most birds fly, but that penguins are birds that do not fly and, learning that Tweety is a bird, infer that it flies. Learning that Tweety is a penguin, will in no way make us change our mind about the fact that most birds fly and that penguins are birds that do not fly, or about the fact that Tweety is a bird. It should make us abandon our conclusion about its flying capabilities, though. It is most probable that intelligent automated systems will have to do the same kind of (nonmonotonic) inferences.

Many researchers have proposed systems that perform such nonmonotonic inferences. The best known are probably: negation as failure [2], circumscription [9], the modal system of [10], default logic [12], autoepistemic logic [11] and inheritance systems [13]. In [6], [5], and [7] (see preliminary versions in [4] and [8]) the first steps towards a general framework in which those many examples could be compared and classified were taken.

In [5], a number of families of nonmonotonic consequence relations were defined. The underlying set of formulas was left quite unspecified, except for the fact that propositional connectives were supposed to be available. In fact the analysis found in [5] and later in [7] is really adequate only for propositional languages. We shall give here preliminary thoughts about the case of first order predicate calculi.

The rational closure of a conditional knowledge base will play a fundamental role in our treatment of predicate calculi. This construction has been proposed in [7] as a reasonable description of the set of conditional assertions entailed by another such set. There, the construction was defined only for finite knowledge bases and given a model-theoretic definition. Since then, this construction, for propositional languages, has been given both an abstract characterization and an algorithmic description. In the same time it has been generalized to arbitrary knowledge bases and appealing global properties of this construction have been shown to hold. This work is currently in progress.

2 Predicate Calculus: Why?

The purpose of this extended abstract is to examine the extension of the authors' previous work, that dealt with propositional languages, to predicate calculus. One may rightly ask whether this is a worthy enterprise. We shall first, therefore, discuss the status of the debate: predicate calculus vs. propositional calculus.

There is no doubt that, among mathematical logicians and especially those interested in the foundations of mathematics, predicate calculus is considered to be the language of choice, richer, more interesting. After all, predicate calculus is the universal language of mathematics and there is no way the full richness of mathematical reasoning may be captured by propositional logic. But, in this respect, the choices of mathematical logicians should not bear too heavily on us. Researchers in Artificial Intelligence have other concerns than studying mathematical reasoning, and all the evidence gathered during this last decade of fruitful research on nonmonotonic reasoning shows that the kinds of reasoning we have to analyze or realize are different in some essential ways from mathematical reasoning.

More to the point is the observation that, without exception, all systems proposed about ten years ago for nonmonotonic reasoning, used predicate calculus as their basic language. All traditional examples in the field are couched in predicate calculus terms, even when, as in the case of the Yale shooting problem for example, they may obviously be translated in propositional terms. One is therefore surprised to notice that almost none of the efforts in nonmonotonic reasoning have been devoted to analyze the role of quantifiers and free variables in nonmonotonic reasoning. One noticeable exception is Adams' [1], but his motivations are quite different. None of the systems proposed have rules to deal with quantifiers: one finds no introduction or elimination rules, the quantifiers simply disappear from the formalism by some magic, free variables appear syntactically but they are implicitely quantified universally. One of the most often used 'magic' is to consider a formula (or default) with free variables as a short-hand for the (possibly) infinite set of formulas obtained by replacing variables by ground terms. But this essentially means that a formula of predicate calculus stands for a set of propositional formulas. Looking at the examples traditional in the field, one is very hard put to find examples dealing with quantifier alternation, with functions, with formulas containing more than one free variable, with defaults whose antecedent and consequent do not contain the same variables. All the problems discussed in the literature on nonmonotonic reasoning may as well be discussed in the framework of propositional logic, and indeed some recent efforts, mainly in the autoepistemic stream, have decided, with good reason, to move to such a propositional framework. Our position is that we find no absolute necessity to move to the predicate framework, that extending our approach to predicate calculus is not at all easy, but that it is probably worthwhile trying, if only to learn more about the propositional case and to understand better what are exactly the problems raised by variables, functions and quantifiers. We shall present here a preliminary report on the state of our efforts.

The major question we are addressing may simply put in the following way: is $\text{Bird}(x) \mid\!\sim \text{Fly}(x)$ a proper way of saying that *birds, normally, fly*?

3 Preferential Reasoning in Predicate calculus

Let L be a first order language, with equality. The greek letters α, β, and so on, will represent arbitrary formulas (not necessarily closed). The letters x, y and so on, variables. If α and β are formulas, then the pair $\alpha \mid\!\sim \beta$ (read *if* α*, normally* β, or β *is a plausible consequence of* α) is called a conditional assertion (assertion in short). The formula α is the antecedent of the assertion, β is its consequent. The meaning we attach to such an assertion, and against which the reader should check the logical systems we shall discuss is the following: if α is true, I am willing to (defeasibly) jump to the conclusion that β is true. In particular, the intuitive meaning of the assertion $\text{Bird}(x) \mid\!\sim \text{Fly}(x)$ is *If x is a bird, it may be sensibly concluded that it flies*, or, more precisely, *normal birds fly.* The reader should notice that we do *not* allow the application of propositional connectives or quantifiers to assertions. The object $\forall x(\text{Bird}(x) \mid\!\sim \text{Fly}(x))$ is not a well-formed syntactic object for us. But, $(\forall x \text{Bird}(x)) \mid\sim (\forall x \text{Fly}(x))$ is an assertion. *Consequence relations* are sets of conditional assertions.

We shall now briefly describe the intended pragmatics. The queries one wants to ask an automated knowledge base are formulas (of L) and query β should be interpreted as: *is β expected to be true*? To answer such a query the knowledge base will apply some inference procedure to the information it has. This information may be divided into two different types. The first type of information consists of a set of conditional assertions describing the soft constraints (e.g. birds normally fly). This set describes what we know about the way the world generally behaves. This set of conditional assertions will be called the knowledge base, and denoted by **K**. The second type of information describes our information about the specific situation at hand (e.g. it is a bird). This information will be represented by a formula, α.

Our inference procedure will work in the following way, to answer query β. It will try to deduce (in a way that is to be discovered yet) the conditional assertion $\alpha \mid\sim \beta$ from the knowledge base **K**. This is a particularly elegant way of looking at the inference process: the inference process deduces conditional assertions from sets of conditional assertions. Clearly any system of nonmonotonic reasoning may be considered in this way.

The following properties of consequence relations have been introduced in [5]. They constitute *preferential reasoning* in the propositional case. In the framework of predicate calculus, the notation $\alpha \models \beta$ has to be understood in the restricted way: in all first order structures, all assignments that satisfy α also satisfy β.

(1) $\alpha \mid\sim \alpha$ **(Reflexivity)**

(2) $$\frac{\models \alpha \leftrightarrow \beta \ , \ \alpha \mid\sim \gamma}{\beta \mid\sim \gamma}$$ **(Left Logical Equivalence)**

(3) $$\frac{\models \alpha \rightarrow \beta \ , \ \gamma \mid\sim \alpha}{\gamma \mid\sim \beta}$$ **(Right Weakening)**

(4) $$\frac{\alpha \mid\sim \beta \ , \ \alpha \mid\sim \gamma}{\alpha \wedge \beta \mid\sim \gamma}$$ **(Cautious Monotony)**

(5) $$\frac{\alpha \mid\sim \beta \ , \ \alpha \mid\sim \gamma}{\alpha \mid\sim \beta \wedge \gamma}$$ **(And)**

(6) $$\frac{\alpha \mathrel{|\!\sim} \gamma \ , \ \beta \mathrel{|\!\sim} \gamma}{\alpha \vee \beta \mathrel{|\!\sim} \gamma} \qquad \textbf{(Or)}$$

Now that we take L to be a first-order predicate calculus, we wish to add the following two rules. They will be discussed and justified below.

(7) $$\frac{\alpha \mathrel{|\!\sim} \beta}{\exists x \alpha \mathrel{|\!\sim} \exists x \beta} \qquad (\exists - \textbf{intr})$$

(8) $$\frac{\exists x \alpha \mathrel{|\!\sim} \beta, \ x \text{ is not free in } \beta}{\alpha \mathrel{|\!\sim} \beta} \qquad (\exists - \textbf{elim})$$

The eight rules above constitute the system **P** for predicate calculus. A consequence relation that satisfies them is said to be preferential. Let us discuss first the $(\exists - \textbf{elim})$ rule, since this will be a short discussion. This rule is a special case of Monotony. Its justification is that if one is ready to jump to the conclusion that β, which does not involve x, is true on the knowledge that there is an element that satisfies α, one should jump to the same conclusion if one learns that x satisfies α since the new information about the value of the variable x does not change in any essential way our conclusions about the world (variables may take any value) as long as these conclusions do not involve x.

Our argument for accepting the rule $(\exists - \textbf{intr})$ is the following. If normal birds fly and if I obtain information to the effect that there is at least one bird in the world, then it is sensible to conclude that there is at least one normal bird in the world and therefore there is at least one flying individual. More generally, we shall see that rule $\exists$ says that *if there is an individual that has property A then it is plausible that there is an individual that has property A and is normal for that property, i.e., there is a typical A*. Notice, indeed, that the following is a derived rule of **P**.

(9) $$\frac{\alpha \mathrel{|\!\sim} \beta}{\exists x \alpha \mathrel{|\!\sim} \exists x \, (\alpha \wedge \beta)}$$

It implies that we shall not be able to consider situations where *normal birds fly*, *normal birds have feathers*, *normal birds have beaks*, *normal birds are not green* but in which we have information to the effect that there are birds but any bird that have beaks, feathers and flies is green. Such information is, for us, contradictory. In other words, we consider impossible situations in which there are birds but there is no typical bird. Such situations have been referred to in the literature as *the lottery paradox* or *the bird-shop paradox*. For us, such situations are indeed paradoxical. Whether this restriction to non-paradoxical situations is bearable is mainly a pragmatical question and only experience will tell. Notice, though, that *default* reasoning (in the largest possible meaning) is mainly useful when the number of individuals is very large, i.e., when we do not expect to have an exhaustive list of all those individuals available. Otherwise, it seems we could get full information and then classical monotonic reasoning is called for. At least, we are aiming at those situations in which there is no exhaustive list of individuals. In such situations, we may as well accept that there is a typical bird, perhaps at the cost of accepting that phantom birds *exist*, in the sense of the quantifier $\exists$. This quantifier would be badly chosen anyway to represent anything like *physical existence.* Should we accept other rules dealing with quantifiers? Two such candidates come to mind. They must be rejected. Let x be an individual variable and t a term, and α^x_t represents the result of replacing x by t in α and is defined only if no variable free in t clashes with a bound variable of α.

(10) $$\frac{\alpha \mathrel{|\!\sim} \beta}{\forall x\alpha \mathrel{|\!\sim} \forall x\beta} \qquad (\forall)$$

(11) $$\frac{\alpha \mathrel{|\!\sim} \beta}{\alpha^x_t \mathrel{|\!\sim} \beta^x_t} \qquad \textbf{(Substitution)}$$

The reasons to reject ($\forall$) are clear. Suppose we think that *birds, normally fly*, written as $\text{Bird}(x) \mathrel{|\!\sim} \text{Fly}(x)$, and we know that *we are talking only about birds*, written as $\forall x \text{Bird}(x)$, we have no reason at all to conclude that *we are talking only about flying things.* The knowledge that everything is a bird has no bearing on our assumption that there may be birds that do not fly.

The reasons to reject **(Substitution)** are much more delicate and they seem to involve two-place predicates in an essential way. Suppose we hold the default that most pets are dogs or cats, or, equivalently, that normal pet owners have dogs or cats as pets. We would like to describe this default as: $\text{Pet}(x, y) \mathrel{|\!\sim} \text{Dog}(y) \vee \text{Cat}(y)$. Suppose also that we know of an individual, John, who likes snakes. We would like to describe this in a default that says that most of John's pets are snakes. The natural way to do that is obviously: $\text{Pet}(\text{John}, y) \mathrel{|\!\sim} \text{Snake}(y)$. If we accepted the rule of **(Substitution)**, we would deduce from the first default: Pet(John, y) $\mid\!\sim$ Dog(y) $\vee$ Cat(y). With the second default, using **(And)** and the fact that the classes of dogs, cats and snakes have an empty intersection, we would deduce: Pet(John, y) $\mid\!\sim$ **false**, which means that it is completely unthinkable that John has a pet. This is obviously unwanted. What is revealed here is that the system **P**, in the propositional case, is such a powerful system that it is incompatible with the very powerful rule of **(Substitution)**. If one considers what happens in other nonmonotonic formalisms, one sees that the only reason **(Substitution)** is accepted by Reiter's Default Logic or McCarthy's Circumscription is that those formalisms are too weak by themselves (i.e., without some additional formalism) to choose between two different extensions: the one in which John's pet is a snake and the one in which it is a dog or a cat. The price Default Logic has to pay is very high: it cannot even conclude that John's pet is a snake, a dog or a cat. Circumscription fares a bit better, but cannot see, without external help, that the second default is more specific than the first one and should therefore preempt it. The fact that **(Substitution)** has to be rejected raises the question of how can we manage without it. How can we show that Tweety flies?

The system **P** is powerful, as was shown in [5], but it does not enable us to conclude that Tweety flies from the information that normal birds fly and that Tweety is a bird. This is the case because preferential reasoning is not capable of inferring the fact that Tweety is a normal bird from the fact that we have no reason to think that it is abnormal. This task must be delegated to the procedure of rational closure, a first limited version of which has been defined in [7]. But, before we discuss technical matters, let us describe on an informal level why Bird(Tweety) $\mid\!\sim$ Fly(Tweety) is derivable from, i.e., in the rational closure of, the knowledge base containing only Bird(x) $\mid\!\sim$ Fly(x). The property of **(Rational Monotony)** described in equation (13) guarantees that, if we have Bird(x) $\mid\!\sim$ Fly(x), we shall also have Bird(x) $\wedge\ x =$ Tweety $\mid\!\sim$ Fly(x), unless we have Bird(x) $\mid\!\sim x \neq$ Tweety. It will be the task of the rational closure operation to make sure that, in the absence of added information, the assertion Bird(x) $\mid\!\sim x \neq$ Tweety does not enter the rational closure. Now, from Bird(x) $\wedge\ x =$ Tweety $\mid\!\sim$ Fly(x), we shall infer, by propositional preferential reasoning that: Bird(x) $\wedge\ x =$ Tweety $\mid\!\sim$ Fly(x) $\wedge\ x =$ Tweety. The intuitive rationale behind this derivation is that completely obvious. Then $\exists x$ (Bird(x) $\wedge\ x =$ Tweety) $\mid\!\sim \exists x$ (Fly(x) $\wedge\ x =$ Tweety) follows by the rule **($\exists$ – intr)**. Replacing antecedent and consequent by logically equivalent formulas, we obtain: Bird(Tweety) $\mid\sim$ Fly(Tweety). This is indeed, we think, an exact description of why we are *right* to think that Tweety flies. In cases where there are some reasons to think that Tweety is not a normal bird (and preferential reasoning is quite good at discovering such situations) then Bird(x) $\mid\sim x \neq$ Tweety will be derivable (preferentially) and therefore in the rational closure. In such a case the assertion Bird(x) $\wedge\, x =$ Tweety $\mid\sim$ Fly(x) will typically not enter the rational closure.

4 Preferential models

We shall now briefly define preferential models (in the predicate calculus case), along the lines of [5], and show that the consequence relations that may be defined by those models are exactly the preferential relations. Similarly for ranked models and rational relations. The semantic restriction that corresponds to the rules ($\exists$ – **intr**) and ($\exists$ – **elim**) are quite natural, though not so easy to manipulate. We adapt the definition of a preferential model found in [5] to predicate calculus. The following definitions are also taken from [5] and justified there.

Preferential models give a model-theoretic account of the way one performs nonmonotonic inferences. The main idea is that the agent has, in his mind, a partial ordering on possible states of the world. State s is less than state t, if, in the agent's mind, s is *preferred* to or more *natural* than t. Now, the agent is willing to conclude β from α, if all *most natural* states that satisfy α also satisfy β.

Some technical definitions are needed. Let U be a set and $\prec$ a strict partial order on U, i.e., a binary relation that is antireflexive and transitive.

Definition 1 *Let $V \subseteq U$. We shall say that $t \in V$ is minimal in V iff there is no $s \in V$, such that $s \prec t$.*

Definition 2 *Let $V \subseteq U$. We shall say that V is smooth iff $\forall t \in V$, either $\exists s$ minimal in V, such that $s \prec t$ or t is itself minimal in V.*

We may now define the family of models we are interested in. The states will be labeled with worlds. A world should give a truth value to each formula, even formulas that are not closed, and therefore will be defined as a pair

$\langle M, f\rangle$ where M is a first order structure and f assigns an element of the domain of M to each variable.

Definition 3 *A* preferential *model W is a triple $\langle S, l, \prec\rangle$ where S is a set, the elements of which will be called states, $l : S \mapsto \mathcal{U}$ assigns a world to each state and $\prec$ is a strict partial order on S satisfying the following two conditions. The first one, the* smoothness condition *is: $\forall \alpha \in L$, the set of states $\widehat{\alpha} \stackrel{\text{def}}{=} \{s \mid s \in S,\ s \models\!\equiv \alpha\}$ is smooth, where $\models\!\equiv$ is defined as $s \models\!\equiv \alpha$ (read s satisfies α) iff $l(s) \models \alpha$. The second one* E *is:*

1. *if a state s, labeled with $\langle M, f\rangle$, is minimal in $\widehat{\exists x \alpha}$, then there exists a state t that is minimal in $\widehat{\alpha}$ and that is labeled with $\langle M, f'\rangle$ where f' differs from f at the most in the element it assigns to x*

2. *if a state t is minimal in $\widehat{\alpha}$ and labeled with $\langle M, f\rangle$, then there is a state s, labeled with $\langle M, f'\rangle$, where f' differs from f at the most in the element it assigns to x, that is minimal in $\widehat{\exists x \alpha}$.*

The model W will be said to be finite if S is finite.

The smoothness condition is only a technical condition needed to deal with infinite sets of formulas, it is always satisfied in any finite preferential model, and in any model in which $\prec$ is well-founded (i.e., no infinite descending chains).

We shall now describe the consequence relation defined by a model.

Definition 4 *Suppose a model $W = \langle S, l, \prec\rangle$ and $\alpha, \beta \in L$ are given. The consequence relation defined by W will be denoted by $\mid\!\sim_W$ and is defined by: $\alpha \mid\!\sim_W \beta$ iff for any s minimal in $\widehat{\alpha}$, $s \models\!\equiv \beta$.*

If $\alpha \mid\!\sim_W \beta$ we shall say that the model W satisfies the conditional assertion $\alpha \mid\!\sim \beta$, or that W is a model of $\alpha \mid\!\sim \beta$.

It is easy to see that any preferential model that satisfies our additional condition (E) defines a preferential consequence relation that satisfies the rules ($\exists$ – **intr**) and ($\exists$ – **elim**).

The representation theorem is the following.

Theorem 1 *Let $\mid\!\sim$ be a preferential relation. There is a preferential model that defines $\mid\!\sim$.*

Proof: The proof parallels the corresponding proof of [5], only the main steps will be sketched. The only difference is that we restrict our attention to a subset of the possible worlds. Let D be an infinite large enough set of constants not included in L. We shall consider first order structures on the extended language $L \cup D$.

Definition 5 *A first order structure M, on the extended language, is said to be satisfactory iff, given any formula α (in the original language) of the form $\exists x \beta$ and any assignment f of elements of the domain of M to the variables, for which the world $w = \langle M, f \rangle$ is normal for α, there is a constant $d \in D$ such that the world $\langle M, f_x^d \rangle$ (again we should have written d_M instead of d) is normal for β. A world $\langle M, f \rangle$ will be termed satisfactory iff M is satisfactory.*

The reader may check that, changing *world* to *satisfactory world* leaves the completeness proofs of [5] and [8] correct if only one can prove the following lemma 1. On the other hand, if a canonical model is built with satisfactory worlds only, then it satisfies (E), if we can prove lemma 2.

Lemma 1 *If $\alpha \not\mid\!\sim \beta$, then there exists a satisfactory world, that is normal for α and does not satisfy β.*

Proof: Let $\Delta \stackrel{\text{def}}{=} \{\neg\beta\} \bigcup \{\gamma \mid \alpha \mid\!\sim \gamma\}$. Clearly Δ is satisfiable (see [5]). Let T_0 be the logical closure of Δ (over the original language). We shall build an ascending sequence of consistent logically closed sets: T_i on larger and larger languages L_i. The languages L_i will contain the original language L and a finite subset of D. We are going to enumerate all pairs consisting of a formula α of the form $\exists x \beta$ (in the original language) and an assignment of elements of $D \cap L_i$ to some of the free variables of α. Suppose we have defined T_i and L_i and are now dealing with $\alpha = \exists x \beta$ and g that assigns an element of $D \cap L_i$ to some of the free variables of α. If γ is a formula (over L) we shall denote by γ_g the formula obtained by replacing those free variables of γ that have an image under g by their image. If there exists a formula γ (in the language L_i) such that $\alpha \mid\!\sim \gamma$ and $\gamma_g \notin T_i$ then choose any one of those γ's and take T_{i+1} to be the logical closure (over L_i) of $T_i \cup \{\neg\gamma_g\}$ and L_{i+1} to be L_i. The set T_{i+1} is clearly consistent. Otherwise, all such γ_g are in T_i. Let then d be an element of D not in L_i. We shall take L_{i+1} to be $L_i \cup \{d\}$ and T_{i+1} to be the logical closure of the set $\Delta_i = T_i \cup \{\left(\eta_x^d\right)_g \mid \beta \mid\!\sim \eta\}$. It is left to show that the set Δ_i is consistent. Suppose not. Then there is a finite subset of Δ_i that is inconsistent and, since $\mid\!\sim$ satisfies **And**, there is some η such that $\beta \mid\!\sim \eta$ and $T_i \models \neg\left(\eta^d_x\right)_g$. From $\beta \mid\!\sim \eta$, by **($\exists$ – intr)** and **(Right Weakening)** (since x is not free in η^d_x), we have $\alpha \mid\!\sim \eta^d_x$. By hypothesis, then, $\left(\eta^d_x\right)_g$ must be in the set T_i. A contradiction to the fact that T_i is consistent. It is clear that, by dovetailing, one can arrange for the enumeration to contain all pairs of existential formulas and partial assignments into D. The set $T_\infty \stackrel{\text{def}}{=} \bigcup_{i=0}^{\infty} T_i$ is clearly consistent. Any world w that is a model for this set is clearly normal for α and does not satisfy β. Let us check it is satisfactory. Let M be the first order structure of w. Suppose $\alpha = \exists x \beta$ is a formula of L and f is an assignment of elements of D to the variables for which $\langle M, f\rangle$ is normal for α. The pair consisting of α and the assignment f restricted to the free variables of α has appeared somewhere in the enumeration. At this point we certainly did not take the first possibility, otherwise there would be a γ such that $\alpha \mid\!\sim \gamma$ and $\neg\gamma_f \in T_\infty$, which implies that $\langle M, f\rangle$ does not satisfy γ and is not normal for α. Therefore, at this point, we chose the second possibility and the new d introduced in the language at this point is such that for every γ such that $\beta \mid\!\sim \gamma$, w satisfies $\left(\gamma^d_x\right)_f$, i.e., $\langle M, f^d_x\rangle$ is normal for β. ▮

Lemma 2 *If a world w $\langle M, f\rangle$ is satisfactory then, given any formula of the form $\exists x \alpha$ for which w is normal, there is an element e in the domain of M such that the world $\langle M, f^e_x\rangle$ is satisfactory and satisfies all formulas β such that $\alpha \mid\!\sim \beta$.*

Proof: Take e to be d_M for the d whose existence is asserted by definition 5. ▮

5 Renaming

In the system presented so far the rule

$$(12) \quad \frac{\alpha(x) \mid\!\sim \beta(x)}{\alpha(y) \mid\!\sim \beta(y)} \qquad \textbf{(Renaming)}$$

when y is not a free variable of α or β, is not a derived rule. One may argue pro and con invariance under renaming of variables. The corresponding semantic restriction is not difficult to describe and the characterization theorem is not difficult either. The question of invariance under renaming seems to be completely orthogonal to the quest for rules dealing with quantifiers.

6 Rational closure

As we have argued above, the system of preferential reasoning in predicate calculus has one main weakness related to predicate calculus (it has other weaknesses that relate to propositional calculus too): *Tweety does not fly.* It is therefore necessary to build an additional layer of reasoning on top of preferential reasoning. The rational closure of a conditional knowledge base, has been proposed in [7] as a reasonable description of the set of conditional assertions entailed by another such set. There, this closure operation was defined in model-theoretic terms and for finite knowledge bases only. We have now a definition of rational closure that is both abstract and general. The idea, that we cannot develop here, is that the rational closure of a knowledge base, if it exists, is its preferred rational extension. *Rational* means that the consequence relation satisfies the following additional rule of **(Rational Monotony)**. A representation theorem for rational relations and ranked models (satisfying the condition (E)) is obtained without too much trouble, following the lines of the corresponding result in the propositional case.

$$(13) \quad \frac{\alpha \wedge \beta \not\mathrel{|\!\sim} \gamma \;,\; \alpha \not\mathrel{|\!\sim} \neg\beta}{\alpha \not\mathrel{|\!\sim} \gamma} \qquad \textbf{(Rational Monotony)}$$

Preferred means least in the following ordering. In the next definition, and in the sequel, $\alpha < \beta$ for K means that $\alpha \vee \beta \mathrel{|\!\sim} \neg\beta$ is in K.

Definition 6 *Let K_0 and K_1 be two rational consequence relations. We shall say that K_0 is preferable to K_1 and write $K_0 \prec K_1$ iff:*

1. *there exists an assertion $\alpha \mathrel{|\!\sim} \beta$ in $K_1 - K_0$ such that for all γ such that $\gamma < \alpha$ for K_0, and for all δ such that $\gamma \mathrel{|\!\sim} \delta$ is in K_0, we also have $\gamma \mathrel{|\!\sim} \delta$ in K_1*

2. *For any γ, δ if $\gamma \mathrel{|\!\sim} \delta$ is in $K_0 - K_1$ there is an assertion $\rho \mathrel{|\!\sim} \eta$ in $K_1 - K_0$ such that $\rho < \gamma$ for K_1.*

The intuitive explanation behind definition 6 is the following. Suppose two agents, who agree on a common knowledge base, are discussing the respective merits of two rational relations K_0 and K_1. A typical attack would be: *your relation contains an assertion* $\alpha \mid\sim \beta$ *that mine does not (and therefore contains unsupported assertions).* A possible defense against such an attack could be: *yes, but your relation contains an assertion* $\gamma \mid\sim \delta$ *that mine does not, and you yourself think that* γ *describes a situation much more usual than the one described by* α. Such a defense much be accepted as valid. Definition 6 exactly says that the proponent of K_0 has an attack that the proponent of K_1 cannot defend against (this is part 1) but that he (i.e., the proponent of K_0) may find a defense against any attack from the proponent of K_1 (this is part 2 of the definition).

The relation $\prec$ among rational relations is, as expected, a strict partial order. We conjecture the following (for predicate calculus):

Conjecture 1 *If* **K** *is a finite knowledge base, it has a rational closure*

The corresponding result for propositional calculus has been proved. We also hope to provide an algorithmic characterization of the rational closure of a finite knowledge base over predicate calculus, similar to the one proposed for the propositional case. This characterization leads, again in the propositional case, to an efficient algorithm computing the rational closure of a finite knowledge base.

7 Discussion and conclusion

We have not shown yet that *Tweety flies* may be deduced from, i.e., is in the rational closure, $\overline{\mathbf{K}}$, of the knowledge base **K** that contains the single assertion $\text{Bird}(x) \mid\sim \text{Fly}(x)$. From the discussion at the end of section 3, we know it is enough to show that the assertion $\text{Bird}(x) \mid\sim x \neq \text{Tweety}$ is not the rational closure of **K** (by our conjecture such a closure exists). To show this, remark that the one-state model in which $x = \text{Tweety}$, Bird(Tweety) and Fly(Tweety) hold defines a rational relation R (since it is ranked) that extends **K**. But it does not contain $\text{Bird}(x) \mid\sim x \neq \text{Tweety}$. By definition 6, then, there must be an assertion $\alpha \mid\sim \beta$ in $R - \overline{\mathbf{K}}$, such that, in R, $\alpha < \text{Bird}(x)$. But this is impossible, since $\text{Bird}(x)$ is satisfied at the lowest level in the model.

References

[1] Ernest W. Adams. The logic of high probabilities. *Journal of Philosophical Logic*, 15:258, 1986.

[2] Keith L. Clark. Negation as failure. In H. Gallaire and J. Minker, editors, *Logics and Data Bases*, pages 293–322. Plenum Press, 1978.

[3] Gerhard Gentzen. *The Collected Papers of Gerhard Gentzen, edited by M. E. Szabo.* North Holland, Amsterdam, 1969.

[4] Sarit Kraus, Daniel Lehmann, and Menachem Magidor. Preferential models and cumulative logics. Technical Report TR 88-15, Leibniz Center for Computer Science, Dept. of Computer Science, Hebrew University, Jerusalem, November 1988.

[5] Sarit Kraus, Daniel Lehmann, and Menachem Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 1990. in press.

[6] Daniel Lehmann. Preferential models and cumulative logics. In Ehud Shapiro, editor, *Fifth Israeli Symposium on Artificial Intelligence, Vision and Pattern Recognition*, pages 365–381, Tel Aviv, Israel, December 1988. Information Processing Association of Israel.

[7] Daniel Lehmann. What does a conditional knowledge base entail? In Ron Brachman and Hector Levesque, editors, *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, Toronto, Canada, May 1989. Morgan Kaufmann.

[8] Daniel Lehmann and Menachem Magidor. Rational logics and their models: a study in cumulative logic. Technical Report TR 88-16, Leibniz Center for Computer Science, Dept. of Computer Science, Hebrew University, Jerusalem, November 1988.

[9] John McCarthy. Circumscription, a form of non monotonic reasoning. *Artificial Intelligence*, 13:27–39, 1980.

[10] Drew McDermott and John Doyle. Non-monotonic logic I. *Artificial Intelligence*, 25:41–72, 1980.

[11] Robert C. Moore. Possible-world semantics for autoepistemic logic. In *Proceedings of AAAI workshop on nomonotonic reasoning*, pages 396–401, New Paltz, 1984.

[12] Raymond Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[13] David S. Touretzky. *The Mathematics of Inheritance Systems*. Research Notes in Artificial Intelligece. Pitman, London — Morgan Kaufmann, Los Altos, 1986.