Research Report AI–1994–08

Attunement to Constraints in Nonmonotonic Reasoning

William C. Rounds
Artificial Intelligence Laboratory
University of Michigan

Guo-Qiang Zhang
Department of Computer Science
University of Georgia

Attunement to Constraints in Nonmonotonic Reasoning*

William C. Rounds
Artificial Intelligence Laboratory
University of Michigan
Ann Arbor, MI 48109 USA
and
Guo-Qiang Zhang
Department of Computer Science
University of Georgia
Athens, Georgia 30602

December 7, 1994

1 Introduction

This paper began as an attempt to use domain theory to throw some light on some of the semantic problems in the area of non-monotonic logic within artificial intelligence. After having thought about some of the issues involved, though, it seemed to us that the examples and methods we use might be of interest in a broader setting. We are offering the results of our investigations in that spirit, hoping that this will be a start on the problem of putting the work in standard AI approaches to non-monotonicity together with current work on information flow (Seligman and J.Barwise 1993.)

We are interested in the subject of default or "normal" inferencing. On the surface of things, this can be exemplified within propositional logic using a *non-monotonic consequence relation* $\alpha \mid\sim \beta$ between sentences α and β. The typical gloss of this is "birds normally fly." So, in terms of (one version of) situation theory, one would take α and β as types, as in the "smoke means fire" paradigm of Devlin (Devlin 1991.) In later work, Barwise (Barwise 1989) suggests that this relation, as a relation between types, would be an *indicating relation*, to be accompanied by a corresponding *signaling relation* $s_1 \rightarrow s_2$, where s_1 is a "site" supporting α, and s_2 is a site supporting β. In later

Imagine for a moment that there are no exceptions to any rules. We can then use a "strict" indicating relation $\alpha \Rightarrow \beta$ to mean that any object of type α is also of type β. (The pair (α, β) is called a constraint.) In this

*Research sponsored by NSF Grant IRI-9120851 and a University of Georgia Faculty Reserach Grant. A very preliminary draft of the paper appeared in the collection of position papers for the First Workshop on Principles and Practices of Constraint Programming, Newport, RI, May 1993, sponsored by ONR.

case, we can take the signaling relation to be the identity on the objects (sites). Then we can have a flow of information about such objects: we can associate the new information that an object s is of type β given that it was of type α. So far this is pretty unexceptional, but the interesting point is that *even when there are exceptions*, we still use the identity signaling relation to gain information about objects. That is, if we know that birds normally fly, and only that Tweety is a bird, then we infer (although possibly incorrectly) that Tweety can fly. The point of default reasoning is that we use such information virtually all the time.

This scenario has of course been extensively studied in artificial intelligence. Barwise has also given a situation-theoretic treatment of it in (Barwise 1989, chapter 5), where the additional idea of *background type* was introduced. The idea there was that a background type N might stand for "the type of birds which are not like penguins, emus, and ostriches", so that conditional on the background type N, the type "bird" would correctly entail the type "fly". This idea was subsequently refined, both by Barwise and by Seligman (Seligman and J.Barwise 1993), into the idea of an information channel. In channel theory, pairs $s1 \rightarrow s2$ are *classified* by some channel as being of type $\alpha \Rightarrow \beta$ or not. If this classification holds, and $s1$ is in fact of type α, then we can infer that $s2$ is of type β. Exceptions can occur here: when, for example, the pair $s1 \rightarrow s2$ fails to be classified either positively or negatively. Cavedon (Cavedon 1994) has given a semantics of conditional logic and a semantics for defaults using these ideas directly.

In this paper, we would like to present another information-theoretic proposal to model the above scenario. It also involves the notion of background, but in a somewhat different way. We think of a background constraint as a *strict* constraint relative to which we add non-strict *default constraints* of the kind mentioned above. The kind of background constraint we have in mind is the strict constraint that penguins are birds, and do not fly.

We work in first-order logic, and we model background constraints as conditional first-order sentences. So "penguins are birds" is just given as the usual universally quantified sentence (Horn sentence) expressing the constraint. We use background constraints to govern partial models of first-order logic. These models are constructed using systems of default rules, as in the *default logic* of Reiter (Reiter 1980), but where Reiter's rules build logical theories, our rules build models. Our approach takes advantage of domain-theoretic notions. A system of default rules is a straightforward generalization of Scott's notion of information system (Scott 1982.) We think of these default systems as "programs" which are created by a reasoning agent to satisfy default constraints.

We show how to treat the standard penguin example in our system, and we give what we hope are some other amusing case studies. The interesting observation here, we feel, is that we can show a specific example of what

Devlin calls “attunement” to background constraints. The idea is that a reasoner will create default systems in response to experience, in an effort to make default constraints into hard ones. But these systems of defaults can be giving a lot of non-information and even false information in a probabilistic sense. They, and the default constraints themselves, should be undergoing revision. We illustrate this by considering an anomaly of Poole (Poole 1989), related to the so-called “lottery paradox” (Jr. 1961); and we consider a more complex case involving the well-known (folk?) “Nixon Diamond.”

To this end, we introduce a notion both of non-monotonic consequence $\mid\!\sim$ which can be used to state default constraints, and, in the finite case, a probabilistic notion of “degree of belief” useful in analyzing the examples.

2 Relations to standard AI approaches

Etherington (Etherington 1988) gave the first model-theoretic semantics to Reiter’s default logic. This was a system based on *sets* of first-order models. Marek, Nerode, and Remmel (Marek et al. 1990), gave a semantics for nonmonotonic rule systems. They translated Reiter’s default rules into finitary formulas of a certain special infinitary logic. Extensions – the central construct of Reiter’s logic – are viewed as models for certain formulas encoding the existence of default derivations.

Our approach has certain commonalities with the Nerode, Marek, and Remmel theory, in that we view extensions model-theoretically. However, we use extensions as models for ordinary first-order logic, not the special logic used by Marek, Nerode and Remmel. It will also be clear that first-order logic is not the only possible logic for which default models could serve as a semantic space. But we concentrate on the first-order case, since that involves the use of constraints.

Our treatment also has the advantage that one can analyze default reasoning situations by working directly with models, as one does all the time in ordinary mathematical reasoning. This contrasts with the approaches of MNR and Etherington, where in the first case, the logic describes a proof theory, and in the second, where one works with sets of first order models as models for default logic. We hold the thesis that Reiter’s default systems should be regarded, not as proof rules, but as *algorithms for the direct construction of partial models for some appropriate logic.* This is a simple and radical reconstruction of default reasoning. To give it a proper explication, though, we use domain-theoretic tools – information systems and Scott domains, in particular, since in our view default reasoning is about what happens when information necessary to resolve a question is lacking.

As we have stated, we want our models to be governed by *constraints*, which in our setting are thought of as laws which govern the behavior of partial models, but which are in the background. We encode a constraint theory into the monotonic forcing relation $\vdash$ of a Scott information system appropriate for a first-order logic semantics. How to accomplish this encoding was not absolutely clear. One possibility is to use a generalization of information systems themselves, due to Saraswat and Panangaden (Saraswat et al. 1991), to the first-order case. We have determined, however, that such a move is unneccessary. We represent constraint theories as a special case of ordinary monotonic information systems.

The next problem is how to add a non-monotonic component to information systems. This we have done by simply adding default forcing rules to Scott's systems.

A final problem is how to use the domains we generate as models for first-order logic itself, and specifically, how to interpret negation. We have chosen a restricted, positive version of first-order logic, which only allows negation on atomic formulas. Then we introduce a notion $\mid\!\sim$ of *non-monotonic consequence* between sentences of first-order logic, as in Kraus, Lehmann, and Magidor (Kraus et al. 1990.) We say that in a default model structure, one sentence non-monotonically entails a second if the second holds in all extensions of "small" partial models of the first. Here "smallness" is interpreted with respect to the natural partial order associated with a Scott domain.

We then turn to the construction of degrees of belief using finite models. Finite default models are of course a special case of our theory. We can generalize the usual finite model theory to partial models, and can use default rules to assign probabilities to statements in FOL, representing an agent's degree of belief in certain situations obtaining. This gives a way of thinking about the usual "birds normally fly" as a pseudo-probabilistic statement. We illustrate this method in the resolution of an anomaly with standard default reasoning, due to Poole (Poole 1989.)

The paper is organized as follows. In section 3 we cover the basics of domain theory and information systems, introduce our non-monotonic generalization, and state a representation theorem for default domains. In Section 4 we show how to interpret first-order positive logic using default models. This is where constraints play a crucial role. Then in Section 5 we introduce our notion of conditional degree of belief, and treat our examples.

3 Default Domain Theory

3.1 Information Systems

We review Scott's representation of domains using *information systems*, which can be thought of as general concrete monotonic "rule systems" for building Scott domains.

Definition 1 An information system is a structure $\underline{A} = (A,\ Con,\ \vdash)$ where

- A is a countable set of tokens,
- $Con \subseteq \mathrm{Fin}\,(A)$, the consistent sets,
- $\vdash \subseteq Con \times A$, the entailment relation,

which satisfy

1. $X \subseteq Y\ \&\ Y \in Con \Rightarrow X \in Con$,
2. $a \in A \Rightarrow \{a\} \in Con$,
3. $X \vdash a\ \&\ X \in Con \Rightarrow X \cup \{a\} \in Con$,
4. $a \in X\ \&\ X \in Con \Rightarrow X \vdash a$,
5. $(\forall b \in Y.\, X \vdash b\ \&\ Y \vdash c) \Rightarrow X \vdash c$.

Example: Approximate real numbers. For tokens, take the set A of pairs of rationals $\langle q, r\rangle$, with $q \leq r$.

The idea is that a pair of rationals stands for the "proposition" that a yet to be determined *real number* is in the interval $[q, r]$ whose endpoints are given by the pair.

Define a finite set X of "intervals" to be in in Con if X is empty, or if the intersection of the "intervals" in X is nonempty. Then say that a set $X \vdash \langle q, r\rangle$ iff the intersection of all "intervals" in X is contained in the interval $[q, r]$. Note that there is only atomic structure to these propositions. We cannot negate them or disjoin them.

The representation of Scott domains uses the auxiliary construct of ideal elements.

Definition 2 An (ideal) element of an information system $\underline{A}$ is a set x of tokens which is

1. consistent: $X \subseteq x \Rightarrow X \in Con$,
2. closed under entailment: $X \subseteq x\ \&\ X \vdash a \Rightarrow a \in x$.

The collection of ideal elements of $\underline{A}$ is written $|\underline{A}|$.

Example. The ideal elements in our approximate real system are in 1-1 correspondence with the collection of closed real intervals $[x, y]$ with $x \leq y$. Although the collection of ideal elements is partially ordered by inclusion, the domain being described − intervals of reals − is partially ordered by reverse interval inclusion. The total or maximal elements in the domain correspond to "perfect" reals $[x, x]$. The bottom element is a special interval $(-\infty, \infty)$.

It can be easily checked that for any information system, the collection of ideal elements ordered by set inclusion forms a Scott domain. Conversely, every Scott domain is isomorphic to a domain of such ideal elements. These results are basic in domain theory, and have been generalized to other classes of complete partial orders by Zhang (Zhang 1991) and others.

3.2 Default Information Systems

We generalize the theory of information systems by simply adding a default component. We should mention at this point that we limit ourselves to the so-called *normal* default structures. The reason for this is not that we cannot define general default rules, but rather that there are problems with the existence of extensions in the full case that we want to avoid.

Definition 3 A normal default information structure is a tuple

$$\underline{A} = (A,\, \mathit{Con}, \Delta,\, \vdash)$$

where $(A, \mathit{Con}, \vdash)$ is an information system, Δ is a set of pairs (X, Y) of consistent finite subsets of A, each element of which is written as $\dfrac{X : Y}{Y}$.

In our application, tokens will be "tuples" or *infons* of the form

$$\langle\!\langle \sigma, m_1, \ldots, m_n; i \rangle\!\rangle,$$

where σ is a relation name, the m_j are elements of a structure, and i is a "polarity" – either 0 or 1. The rules in Δ should therefore be read as follows. If the set of tuples X is in our current database, and if adding Y would not violate any database constraints, then add Y.

In default logic, the main concept is the idea of an extension. We define extensions in default model theory using Reiter's conditions, but extensions are now (partial) models. The following definition is just a reformulation, in information-theoretic terms, of Reiter's own notion of extension in default logic.

Definition 4 Let $\underline{A} = (A, \Delta, \vdash)$ be a default information structure, and x a member of $|\underline{A}|$. For any subset S, define $\Phi(x, S)$ to be the union $\bigcup_{i \in \omega} \phi(x, S, i)$, where

$$\phi(x, S, 0) = x,$$
$$\phi(x, S, i+1) = \overline{\phi(x, S, i)} \cup \bigcup \{Y \mid \frac{X : Y}{Y} \in \Delta \;\&\; X \subseteq \phi(x, S, i) \;\&\; Y \cup S \in \mathit{Con}\}.$$

y is an extension of x if $\Phi(x, y) = y$. In this case we also write $x \delta_{\underline{A}} y$, with the subscript omitted from time to time.

Example: Default approximate reals. Use the information system described above. We might like to say that "by default, a real number is either between 0 and 1, or is the number π". We could express this by letting Δ consist of the rules $\dfrac{: Y}{Y}$, where Y ranges over singleton sets of rational pairs $\{\langle p, q \rangle\}$ such that $p \le 0$ and $q \ge 1$, together with those pairs $\{\langle r, s \rangle\}$ such that $r < \pi$ and $s > \pi$. Then, in the ideal domain, the only extension of $[-1, 2]$ would be $[0, 1]$; the interval $[-2, 0.5]$ would have $[0, 0.5]$ as an extension, and there would be 2 extensions of $[-2, 4]$, namely $[0, 1]$ and $[\pi, \pi]$.

In the full paper we show that all of this material can be stated in order-theoretic terms, without the need for information systems. This will make it possible to see the essential formula defining extensions, and will give a hint as to why we believe the order-theoretic approach is an interesting one to take.

4 Constraint default structures for first-order logic

Assume, for purposes of this paper, that we are given a signature for first-order logic without equality, and with no function symbols other than constants. We will interpret first order logic using a nonstandard class of models. Our structures will be default information systems based on a particular set of individuals M. We first have to assume some *constraints* on any relations which are going to be holding in such sets M. These constraints will be used to generate the monotonic forcing relation $\vdash$ in the default structure. (The defaults themselves can be arbitrary, as long as they are normal.) We can use sets C of arbitrary closed formulas of first-order logic to state background constraints; in fact, we can use any language for which first-order structures are appropriate models.

To interpret formulas, we first of all choose some set M of individuals. We do *not* fix relations on M as in the standard first-order case, but we do choose particular individuals to interpret the constants[1]. Now, tokens will be infons of the form

$$\sigma = \langle\!\langle R, m_1, \ldots, m_n; i \rangle\!\rangle$$

where R is a relation name, $m_j \in M$, and $i \in \{0, 1\}$. (This last item is called the *polarity* of the token.) We say that a set s of these tokens is *admissible* if (i) it does not contain any tokens conflicting in polarity, and (ii) it matches a model of C in the usual first-order sense. That is, there is a structure

$$\mathcal{M} = (M, (R_1, \ldots R_k))$$

where the R_j are relations on M of the appropriate arities, such that $\mathcal{M}$ is a model of C, and such that

$$\langle\!\langle R_j, m_1, \ldots, m_n; 1 \rangle\!\rangle \in s \Rightarrow R_j(m_1, \ldots, m_n) \text{ is true in } \mathcal{M}.$$

Similarly,

$$\langle\!\langle R_j, m_1, \ldots, m_n; 0 \rangle\!\rangle \in s \Rightarrow R_j(m_1, \ldots, m_n) \text{ is false.}$$

An admissible set of infons is *total* if it is maximal in the subset ordering on sets of infons. A total set is isomorphic to an ordinary first-order structure $\mathcal{M}$.

Now we can specify a default information structure relative to M and C. Actually, the work is in specifying the strict (monotonic) part of the system. The defaults can be arbitrary normal ones.

[1] In terms of philosophy of language, we are taking constants to be rigid designators.

Definition 5 Let M be a set, and C a constraint set. A first-order default information structure relative to M and C is a structure of the form

$$\underline{A}(M, C) = (\, A,\ Con,\ \Delta,\ \vdash\,)$$

where A is the token set described above. A finite set X of tokens will be in *Con* if it is admissible, and $X \vdash \sigma$ iff for any total admissible set t, if $X \subseteq t$ then $\sigma \in t$.

Examples. The above definition encodes the constraints C into the $\vdash$ relation of the information system. For example, consider the constraint obtained by taking C to be the true formula $\mathbf{t}$. Intuitively, this should be no constraint at all, so our entailment relation should be the minimal one in which $X \vdash \sigma$ if and only if $\sigma \in X$. This is in fact the case. First notice that because $C = \mathbf{t}$, that a total admissible set t is one which (i) contains no infon $\sigma = \langle\!\langle \sigma, m; i \rangle\!\rangle$ and the *dual* infon $\overline{\sigma}$ of opposite polarity; and (ii) for any infon σ, contains either σ or $\overline{\sigma}$. Now let X be a finite set of infons. If $X \vdash \sigma$ then by properties of information systems, the dual infon $\overline{\sigma} \notin X$. By definition of $\vdash$, for any total admissible set t of infons, if $X \subseteq t$ then $\sigma \in t$. If σ is not in X, let t be a total admissible set containing X and the infon $\overline{\sigma}$ of opposite polarity. Then both σ and $\overline{\sigma}$ would be in t, which is not possible for an admissible set.

Notice that our general definition is easily modified to particular classes of interpretations. For example, our constraints may be stated for just one intended model, say the real numbers with addition and multiplication. In that case, we choose our sets M to be allowable by the particular interpretation class, and we change the definition of admissibility so that first-order structures are chosen from our particular class as well. Technically, we should restrict M to be countable, so that our Scott domain is in fact ω-algebraic. In fact, though, we will mostly be interested in *finite* default models for first-order logic.

4.1 Syntax and Semantics

For lack of space, we omit the official details of our three-valued semantics; but they are standard, given a knowledge of the strong Kleene truth tables.

4.2 Nonmonotonic consequence

Our semantics can now be used to define a relation of nonmonotonic entailment, written $\mid\!\sim$, between sentences of our (positive) first-order logic. Understanding this notion is a step towards understanding the probabilistic measure introduced in the next subsection.

Intuitively, when we say that φ nonmonotonically entails ψ, we mean that having only the information φ, we can "leap to the conclusion" ψ. The usual example is that, knowing only that Tweety is a bird, we can leap to the conclusion that Tweety flies, even though penguins do not fly. A great deal of effort in the AI community has gone into giving a proper interpretation to the assertion $\varphi \mid\!\sim \psi$. We use (finite) default models and extensions to interpret it.

The notion of "only knowing" φ in $\varphi \mid\!\sim \psi$ Levesque 1990, given a default information structure, is captured by interpreting the antecedent φ in a certain small class of situations for the structure. There are at least two possibilities for this class. One natural one is to use all set-theoretically minimal situations supporting φ. The second is to interpret φ in the *supremum closure* of these minimal models. We choose the second in this paper, because it seems better motivated from the probabilistic standpoint to be given in the next subsection.

We therefore make the following definitions.

Definition 6 Let $\underline{A}(M, C)$ be a default information structure, and φ a sentence of our logic. Let s, t range over situations.

- $MM(\varphi)$ is the set $\{s \mid s \text{ is minimal such that } s \models \varphi\}$;
- $U(\varphi)$ is the **supremum closure** of $MM(\varphi)$: the collection of situations obtained by taking consistent least upper bounds of arbitrary subcollections of $MM(\varphi)$. If $s \in U(\varphi)$ we will say that s is a minimal-closure model of φ.

Notice that since our logic is positive, every situation in $U(\varphi)$ will support φ.

Given these concepts, we can define nonmonotonic consequence (in a structure) as follows.

Definition 7 Let φ and β be sentences in first-order logic. Let $\underline{A} = A(M, C)$ be a finite normal default information system as above. We say that $\varphi \mid\!\sim_{\underline{A}} \beta$ if for all minimal-closure models $s \in U(\varphi)$,

$$\forall t: \ t \text{ is an } \underline{A}\text{-extension of } s \ \Rightarrow t \models \beta.$$

Example. We give the standard bird-penguin example. Assume that our language contains two predicates $Bird$ and $Penguin$, and that $Tweety$ is a constant. Let C be the constraint

$$(\forall x)(Penguin(x) \rightarrow Bird(x) \wedge \neg Fly(x)).$$

Consider a structure $A(M, C)$. Form the defaults

$$\frac{\langle\langle bird, m; 1\rangle\rangle : \langle\langle fly, m; 1\rangle\rangle}{\langle\langle fly, m; 1\rangle\rangle}$$

for each m in M. These defaults express the rule that birds normally fly. We then have

$$MM(Bird(Tweety)) = U(Bird(Tweety)) = \{\{\langle\langle Bird, tw; 1\rangle\rangle\}\}$$

where tw is the element of M interpreting $Tweety$.

The only extension of $\{\langle\langle Bird, tw, ; 1\rangle\rangle\}$ is $\{\langle\langle bird, tw; 1\rangle\rangle, \langle\langle fly, tw; 1\rangle\rangle\}$. Therefore

$$Bird(Tweety) \mid\!\sim Fly(Tweety).$$

We do not have $Penguin(Tweety) \mid\sim Fly(Tweety)$, because of the constraint C.

5 Probabilities, Constraints, and Attunement

Where do the default structures (in particular the default systems above) come from? We suggest that they could come from *default constraints*. Consider a (syntactic) construct of the form $\phi(x) \mid\sim \psi(x)$ where x is a free variable (perhaps *parameter*). Then, given a structure $A = A(M, C)$ an *admissible default system* would be one where the set of defaults Δ is such that with respect to all anchorings α of the free variables, we have that $(\phi \mid\sim_A \psi)[\alpha]$. By this last notation, we mean just to write out the definition of $\mid\sim$ again, but with respect to the anchoring α. A more stringent notion of consequence is now possible, as we can insist that in a structure, one formula entails another with respect to any admissible default system.

In fact, though, the usual default sets seem to come about in other ways. The example of the Nixon Diamond will serve to illustrate this point. In this example, Quakers are by default pacifists, and Republicans by default warmongers, and Nixon is strictly a Quaker and a Republican. The default sets satisfying the two default constraints are usually lumped together, with the result that one extension has Nixon as a warmonger, and another has him as a pacifist. No one so far has tried to separate default sets, constructing extensions in stages, to our knowledge.

The fact that default sets can be arbitrary has other amusing ramifications. We can use defaults to generate *degrees of belief* or *subjective probabilities* of various logical statements. By "subjective probability" we mean an analogue of the usual probability, a number that would be assigned to a statement by a particular agent or subject, given a default system and some basic constraints on the world. Let us illustrate with an example of Poole (Poole 1989.)

5.1 Poole's anomaly

Assume that there are exactly three mutually exclusive types of birds: penguins, hummingbirds, and sandpipers. It is known that penguins don't fly, that hummingbirds are not big, and that sandpipers don't nest in trees. Now suppose we want to assert that the typical bird flies. Since we only speak of birds, we can do this with the precondition-free "open default constraint"

$$\mathbf{true} \mid\sim fly(x).$$

We would also like to say that the typical bird is (fairly) big, and that it nests in a tree. Similar defaults are constructed to express these beliefs.

The " paradox" is that it is impossible now to believe that Tweety, the typical bird, flies. To see why, let us formalize the problem more fully in our first-order language. Let C_1 be the obvious first-order sentence asserting that every individual is one of the three types of birds, and that no individual is of more than one type. Let C_2 be the conjunction of three sentences expressing abnormalities. One of these is, for example,

$$\forall x(\mathit{Penguin(x)} \rightarrow \neg \mathit{fly(x)}).$$

Let the background constraint C be $C_1 \wedge C_2$.

The defaults must be given in the semantics. Let M be a finite set, and consider the first-order default structure A with the admissible set of precondition-free defaults

$$\frac{:\langle\!\langle \mathit{fly}, m; 1\rangle\!\rangle}{\langle\!\langle \mathit{fly}, m; 1\rangle\!\rangle}; \frac{:\langle\!\langle \mathit{big}, m; 1\rangle\!\rangle}{\langle\!\langle \mathit{big}, m; 1\rangle\!\rangle}; \frac{:\langle\!\langle \mathit{tree}, m; 1\rangle\!\rangle}{\langle\!\langle \mathit{tree}, m; 1\rangle\!\rangle}.$$

We need only precondition-free defaults, because we only speak about birds. Further, we need no infons mentioning penguins, or any of the other species. The constraints can still operate.

We assert that if M has n elements, then there are 3^n extensions of the empty set (which is in fact the least model of $\mathbf{t}$). This is because any extension will include, for each bird $m \in M$, exactly two out of the three infons $\langle\!\langle fly, m; 1\rangle\!\rangle, \langle\!\langle big, m; 1\rangle\!\rangle, \langle\!\langle tree, m; 1\rangle\!\rangle$. The extension cannot contain all three infons, because the constraints rule that out. So each of n birds has three choices, leading to 3^n extensions.

One such extension is

$$\{\langle\!\langle \mathit{big}, m; 1\rangle\!\rangle : m \in M\} \cup \{\langle\!\langle \mathit{tree}, m; 1\rangle\!\rangle : m \in M\}$$

which omits any infons of the form $\langle\!\langle \mathit{fly}, m; 1\rangle\!\rangle$. This extension is one where *no* birds fly, where all birds are penguin-like. So now if *Tweety* is a constant of our language, then the formula *Fly(Tweety)* is not a nonmonotonic consequence of the "true" formula **true**, whose minimal model is the empty set. Further, if we move to the situation of seventeen bird types, each with its own distinguishing feature, we still have the case that Tweety cannot be believed to be flying. Poole suggests that this raises a problem for most default reasoning systems.

5.2 A pseudo-probabilistic solution

We now contend that the problem is not so severe. Notice that it is only in 3^{n-1} extensions that Tweety does not fly. This is because in an extension where Tweety does not fly, the constraints force the infons involving Tweety to assert that he is big and lives in a tree. Tweety thus only has one non-flying choice. The other $n-1$ birds have the same three choices as before. It seems therefore truthful to say that with probability $(3^n - 3^{n-1})/3^n = 2/3$, Tweety believably does fly. Moreover, imagine a scenario with seventeen mutually exclusive bird types, the same kinds of exceptions for each type, and defaults for all of the types. Then we would get that Tweety flies with probability 16/17.

We use this example to define our notion of subjective degree of belief:

Definition 8 Let φ be a sentence of positive first-order logic. Assume that relative to a constraint C, φ has minimal models in a structure M with n elements. Then the **conditional subjective degree of belief** $Pr([\psi \mid \varphi]; n, C)$ is defined to be the quantity

$$\frac{1}{N} \sum_{s \in U(\varphi)} \frac{card\{e : s\ \delta\ e\ \&\ e \models \psi\}}{card\{e : s\ \delta\ e\}}$$

where N is the cardinality of $U(\varphi)$.

Example. Referring to the case of Tweety above, we have

$$Pr([Fly(Tweety) \mid \mathbf{t}]; n, C) = 2/3.$$

Note the non-dependence on n. This raises a question about limits as n gets large, a topic which we must defer here. The full story is more nearly told in (Grove et al. 1992), which refers to the notion of *limit law* for models in first-order logic. The problem seems to be that in many examples of the above type, there is covergence of the subjective degree of belief measure in models of size n as n grows without bound. However, there seems to be no characterization of exactly when this happens, and simple (non-natural) examples show that limits need not always exist.

Here is what we mean now by "attunement." Notice that if we change our set of predicates and constraints to the case of seventeen bird types, but retain the rule system for three types, then we still get the same degree of belief (2/3) for Tweety's flying. Imagine that the universe had had seventeen bird types all along, with the constraints for those types. Then our agent, living in a small portion of the world (is there a situation with penguins and hummingbirds in it?) might have only observed the three types of birds. In that case, her 2/3 subjective degree of belief would not be as correct as it could be. Traveling to Australia might help refine the defaults.

Our definition bears a strong resemblance to the notion defined by Bacchus, Grove, Halpern, and Koller (Bacchus et al. 1993.) Their definition of the conditional probability of a statement ψ given another statement φ, though, is not made with reference to a given default information system. Instead, defaults are "translated" into a special logic for reasoning about statistical information. (For example, one can say that the proportion of flying birds out of all birds is approximately .9). Then, the translated default statements, and the given formulas φ and ψ are given a conditional probability in a standard first-order structure. Our corresponding "translation" of default statements is into a system of default rules, just as in Reiter's formulation. Our semantics also contrasts with that of BGHK in that it looks at partial worlds as well as total ones, and can assign degrees of belief to a statement's not being resolved one way or another.

5.2.1 Nixon revisited

The method of default model theory can be adapted to differing kinds of logics for reasoning about default models. This will help us make use of a more specific logical language, should that be appropriate. We illustrate this with an improved model of the Nixon Diamond, using the BGHK probabilistic language.

Example (Nixon.) A fraction α of quakers are pacifists, and a fraction β of republicans are non-pacifists. These are our constraints on any actual world, but which people are pacifists currently is not known. Our version of the logic of BGHK has as atomic formulas

$$\rho x.\varphi(x) \approx_i \alpha,$$

which means that "the proportion of elements x satisfying $\varphi(x)$ is approximately α." Here α is a rational fraction in $[0, 1]$, and the subscript i refers to the i-th component of a "tolerance vector" of positive reals

$$\tau = \langle \tau_1, \ldots, \tau_i, \ldots \rangle$$

which is supplied with a standard finite first-order structure M. The semantics is that $(M, \tau) \models \rho x.\varphi(x) \approx_i \alpha$ if the fraction of domain elements satisfying φ is within τ_i of α. Here we set $i = 1$ and can actually fix $\tau_i = 0$. We thus want our background constraints to be

$$C(\alpha, \beta) = \rho x.(Pac(x) \mid Qu(x)) \approx \alpha \wedge \rho x.(\neg Pac(x) \mid Rep(x)) \approx \beta.$$

This formula uses *conditional expressions* of the form $\rho x.(\psi \mid \theta)$, the semantics of which in BGHK are a bit tricky when there are no domain elements satisfying θ, but which are not a problem in our case, as the expression denotes the fraction of domain elements satisfying $\psi \wedge \theta$ divided by the fraction satisfying θ, and we will always have positive numbers in the denominator.

We are interested in what happens as we vary α and β. But we keep these parameters fixed for what follows. Suppose now that our given information is "there are exactly N real quakers" and "there are exactly M real republicans.", and that there is exactly one quaker-republican, and that Nixon is that one. The Bacchus logic cannot easily express such statements. So instead of calculating a conditional "probability", we just consider a world which has exactly this information. We further simplify matters by assuming that N and M are chosen so that the numbers αN and βM are whole numbers. We consider a model of size $N + M - 1$. What is the degree of belief in Nixon's being a definite, true pacifist? We assume our model consists of the integers from 1 to $N + M - 1$ and interpret Nixon as N.

One world satisfying our conditions is a situation s_0 containing the infons

$$\langle\langle qu, n; 1 \rangle\rangle \text{ for } 1 \leq n \leq N$$

and

$$\langle\langle rep, m; 1\rangle\rangle \text{ for } N \leq m \leq N+M-1.$$

Also we have N is Nixon. Now any permutation of the set of $N+M-1$ elements keeping Nixon fixed will do as another minimal situation expressing that there are exactly N quakers and M republicans. These would all assigned equal probability, and the computations would be the same in each case. Thus to get the subjective degree of belief of Nixon being a true pacifist, it suffices to consider the situation first described, and to calculate the fraction of extensions of this situation in which Nixon is actually a true pacifist. This is the degree to which $Pac(N)$ is believed, or the subjective probability of $Pac(N)$.

In this setting, we are interested in the probability of a formula's not being supported one way or another. So by $\neg\phi$ we will now mean the *weak* negation of ϕ. A situation will support $\neg\phi$ iff it does not support ϕ. We do want to talk about true warmongers, and we will do this with new predicate symbols. The predicate $NPac(x)$ will be interpreted as true warmongering. We therefore have to add the constraint that no individual is a true warmonger and a true pacifist at the same time.

Our infons will have the form $\langle\langle \sigma, m; i\rangle\rangle$, where σ is one of $\{rep, qu, pac\}$. (The predicate symbol $NPac$ will be interpreted by $\langle\langle pac, m; 0\rangle\rangle$.)

We could calculate degrees of belief (1) when there are no defaults, (2), when we have the default constraint only that typically republicans are warmongers, (3) when we have only the default constraint that typically quakers are pacifists, and (4) when we have both (2) and (3). The defaults satisfying these constraints are taken to be

$$\frac{\langle\langle rep, m; 1\rangle\rangle : \langle\langle pac, m; 0\rangle\rangle}{\langle\langle pac, m; 0\rangle\rangle} \; (N \leq m \leq N+M+1);$$

and

$$\frac{\langle\langle qu, n; 1\rangle\rangle : \langle\langle pac, n; 1\rangle\rangle}{\langle\langle pac, n; 1\rangle\rangle} \; (1 \leq n \leq N).$$

The first case is easily handled. There is only one extension, namely the current world. In this world, true pacifism is not known, and true warmongering is not known. So "neither", namely $\neg Pac(N) \wedge \neg NPac(N)$ has probability 1. $Pac(N)$ and $NPac(N)$ both have probability 0.

We omit cases (2) and (3) for lack of space, and proceed to case (4). A detailed calculation reveals

$$Pr(Pac(N) \mid s_0) = \frac{\alpha(1-\beta)}{1-\alpha\beta}.$$

Similarly, the degree of belief in Nixon's warmongering is

$$Pr(NPac(N) \mid s_0) = \frac{\beta(1-\alpha)}{1-\alpha\beta}.$$

The degree of belief in neither (the agnostic position) is

$$Pr(\neg(Pac(N) \vee NPac(N)) \mid s_0) = \frac{1-(\alpha+\beta)+\alpha\beta}{1-\alpha\beta}.$$

In all of these expressions s_0 is just our starting situation.

Notice how the choice of α and β influences the number $\frac{\alpha(1-\beta)}{1-\alpha\beta}$ If they are chosen close to 1, the value is indeterminate, unless assumptions are made about how α and β approach 1. When $\alpha = \beta$, for example, then we get a .5 limit. We also get a .5 limit for warmongering, and an 0 limit for being undecided. On the other hand, when $\alpha = \beta = 0$, we get all of the weight on being undecided. But this is natural given that our defaults "program" us to assume that quakers are normally pacifists, and republicans are normally warmongers, in the face of the background constraint that there are no quaker pacifists, nor any republican warmongers. In this case we are stuck with our initial situation. Again we see a case of attunement. If our default constraints are unrealistic, we cannot use them to get information.

References

Bacchus, F., A. Grove, J. Halpern, and D. Koller. 1993. Statistical Foundations for Default Reasoning. In *Proceedings of IJCAI*.

Barwise, Jon. 1989. *The Situation in Logic*. 17. Stanford, California: Center for Study of Language and Information.

Cavedon, Lawrence. 1994. Constraints, Channels, Conditionals. This proceedings.

Devlin, K. J. 1991. *Logic and Information*. Cambridge University Press.

Etherington, D. W. 1988. *Reasoning with incomplete information*. Research Notes in Artificial Intelligence. Morgan Kaufman.

Grove, A., J. Halpern, and D. Koller. 1992. Asymptotic conditional probabilities for first-order logic. In *Proc. 24th ACM Symp. on Theory of Computing*, 294–305.

Jr., H. Kyburg. 1961. *Probability and the Logic of Rational Belief*. Wesleyan University Press.

Kraus, S., D. Lehmann, and M. Magidor. 1990. Nonmonotonic reasoning, preferential models, and cumulative logics. *Artificial Intelligence* 44:167–207.

Levesque, H. J. 1990. All I Know: A study in autoepistemic logic. *Artificial Intelligence* 42:263–309.

Marek, W., A. Nerode, and J. Remmel. 1990. A Theory of Nonmonotonic Rule Systems. In *Proceedings of 5th IEEE Symposium on Logic in Computer Science*, 79–94.

Poole, D. L. 1989. What the Lottery Paradox Tells us about Default Reasoning. In *Proceedings of First Annual Conference on Knowledge Representation*. Morgan Kaufmann.

Reiter, Raymond. 1980. A Logic for Default Reasoning. *Artificial Intelligence* 13:81–132.

Saraswat, V., M. Rinard, and P. Panangaden. 1991. Semantic foundations of concurrent constraint programming. In *Proceedings of POPL 1991*.

Scott, Dana S. 1982. Domains for Denotational Semantics. In *Lecture Notes in Computer Science 140*.

Seligman, J., and J.Barwise. 1993. Channel theory: towards a mathematics of imperfect information flow. June.

Zhang, Guo-Qiang. 1991. *Logic of Domains*. Boston: Birkhauser.