Annals of Mathematics and Artificial Intelligence 20 (1997) 195–226

Logical considerations on default semantics[1]

William C. Rounds

Artificial Intelligence Laboratory, University of Michigan, Ann Arbor, Michigan 48109, USA
E-mail: rounds@engin.umich.edu

Guo-Qiang Zhang

Department of Computer Science, University of Georgia, Athens, Georgia 30602, USA
E-mail: gqz@cs.uga.edu

We consider a reinterpretation of the rules of default logic. We make Reiter's default rules into a constructive method of building *models*, not theories. To allow reasoning in first-order systems, we equip standard first-order logic with a (new) Kleene 3-valued partial model semantics. Then, using our methodology, we add defaults to this semantic system. The result is that our logic is an ordinary monotonic one, but its semantics is now nonmonotonic. Reiter's *extensions* now appear in the semantics, not in the syntax. As an application, we show that this semantics gives a partial solution to the conceptual problems with open defaults pointed out by Lifschitz [17], and Baader and Hollunder [2]. The solution is not complete, chiefly because in making the defaults model-theoretic, we can only add conjunctive information to our models. This is in contrast to default theories, where extensions can contain disjunctive formulas, and therefore disjunctive information. Our proposal to treat the problem of open defaults uses a semantic notion of nonmonotonic entailment for our logic, related to the idea of "only knowing". Our notion is "only having information" given by a formula. We discuss the differences between this and "minimal-knowledge" ideas. Finally, we consider the Kraus–Lehmann–Magidor [14] axioms for preferential consequence relations. We find that our consequence relation satisfies the most basic of the laws, and the Or law, but it does not satisfy the law of Cut, nor the law of Cautious Monotony. We give intuitive examples using our system, on the other hand, which on the surface seem to violate these two laws. We make some comparisons, using our examples, to probabilistic interpretations for which these laws are true, and we compare our models to the cumulative models of Kraus, Lehmann, and Magidor. We also show sufficient conditions for the laws to hold. These involve limiting the use of disjunction in our formulas in one way or another. We show how to make use of the theory of complete partially ordered sets, or domain theory. We can augment any Scott domain with a default set. We state a version of Reiter's extension operator on arbitrary domains as well. This version makes clear the basic order-theoretic nature of Reiter's definitions. A three-variable function is involved. Finding extensions corresponds to taking fixed points twice, with respect to two of these variables. In the special case of *precondition-free* defaults, a general relation on Scott domains induced

[1] Version presented at the Third International Symposium on Artificial Intelligence and Mathematics, Fort Lauderdale, Fla, January 1994. Research supported by NSF grant IRI-9120851.

from the set of defaults is shown to characterize extensions. We show how a general notion of domain theory, the logic induced from the *Scott topology* on a domain, guides us to a correct notion of "affirmable sentence" in a specific case such as our first-order systems. We also prove our consequence laws in such a way that they hold not only in first-order systems, but in any logic derived from the Scott topology on an arbitrary domain.

1. Introduction

Since the introduction of nonmonotonic logic (by McDermott, Doyle, Reiter, McCarthy, Moore, and others), much effort has been spent on finding semantical systems for differing forms of nonmonotonic logic. There is by now considerable agreement that models for a nonmonotonic logic should be ordered by some criterion of "preference", so that formulas can be valid in all "preferred" models instead of all models. That way one can believe in the validity of a formula, though it is not in fact valid. This way of unifying the semantics of differing logics was initiated by Shoham [28].

One logic, for a long time, resisted attempts to provide it with a preferential model semantics: Reiter's default logic [22]. Etherington [9] gave a preferential semantics, involving *sets* of first-order models rather than simple models. Other model theories by now are known. Many of these come from other nonmonotonic logics by means of a translation of default theories to the other logics. The article [26] gives a good proposal, as well as an overview of other proposals, including Lin and Shoham [19] and Lifschitz [18].

Our approach differs from all of the above proposals. We regard Reiter's default systems as *semantics*, not syntactic notions. Our preference relation is purely information-theoretic: the most preferred partial models of a sentence are the ones containing the minimal information affirming it. The central concept of Reiter's system – that of an *extension* – is viewed as a way to add default information, via a nonmonotonic fixed-point operator, to one of these minimal partial models. This is a simple but radical reconstruction of default reasoning. We hope that it will prove to be an interesting and profitable one.

Having introduced default systems as semantic notions, we are free to adopt any of a number of ordinary logics as ways to reason about the semantics generated by default structures. The best analogy may be to the semantics of programming languages. Scott domain theory is one way to assign meanings (denotations) to programs. In analogy, we use Scott domain theory to assign meanings to systems of defaults. In the theory of programming, logics such as Hoare logic, dynamic logic, and temporal logic are used to reason about program properties. In analogy, we could consider first-order logic, modal logic, or some other language to reason about systems of defaults.[2]

[2] Notice that this gives us a possibility to *combine* reasoning about programs with reasoning about defaults, because we use domain theory as a common semantic base.

In the paper, after having stated a domain-theoretic version of defaults, we apply our notion to the case of first-order logic. We equip ordinary first-order logic with partial models, also called *situations*. We interpret negation in the strong Kleene three-valued sense, as in the work of Doherty and Lukaszewicz [8] and others. A new aspect of our construction is that we can have partial models which respect background *constraints*. It is then straightforward to add defaults to these partial models.

One application of this new semantics is to the problem of correctly interpreting open defaults. Lifschitz in [17], and Baader and Hollunder in [2], have shown that Reiter's treatment leads to conceptual difficulties. To this end, we introduce a notion of nonmonotonic consequence, as in the Kraus et al. [14] postulates for a cumulative nonmonotonic logic. With respect to a given structure, we say that a formula φ nonmonotonically entails a formula β iff β is positively satisfied in every default extension of any minimal positive model of φ. We show how to treat the problem of open defaults using our approach. The result seems to give a satisfactory treatment of examples like the Baader–Hollunder ones. However, we view this as a preliminary step in the treatment of such phenomena.

Several points should be made about our definition of nonmonotonic consequence. First, it uses extensions as models, not theories. Second, the intuitive interpretation of our entailment is "if I can confirm only the information given by φ, then I can believe (skeptically) β". This idea is related to, but is not the same as, the notion of "only knowing" [11, 16]. We try to capture it by the use of minimal models, which, of course, we must define. The word "positive" is occasioned by our three-valued Kleene semantics. Finally, the use of default models is in some sense anticipated by the work of Guerreiro and Casanova [10] and Lifschitz [17]. Out definition is much more radical than theirs. Although their semantics involves a fixed-point construction is model-theoretic space, they still regard extensions as theories. Lifschitz in particular also uses a "fixed-universe" construction to deal with extension for open defaults, as we do. However, our extensions are directly constructed by default rules operating in semantic space, whereas for Lifschitz, extensions are only to be found in the syntactic domain.

Another comparison to the set of "minimal-knowledge" approaches mentioned, for example, by Schwarz and Truszczynski, is in order. For these approaches, knowledge is measured by sets of formulas, or theories. One starts with a given set I of formulas. In an intuitive sense, one wants to capture the fact that the theory I is the only theory that is known. To do this, one looks at potential models M of I and chooses *inclusion-maximal* such. Since $M \subseteq N$ implies $Th(N) \subseteq Th(M)$, we see that maximizing models implies minimizing knowledge. (We note here that one speaks about modal formulas, for which models are sets of propositional valuations, so that it makes sense to compare models by inclusion.) The precise details of the differences between the model constructions of Schwarz and Truszczynski, and those of other authors, such as Halpern and Moses [11], need not concern us here, because we are treating a differing notion: that of the minimal *information* conveyed by a formula or theory. This information is measured by the inclusion relation on partial models directly, and does not contain any kind of self-knowledge. For example, the minimal information conveyed by the formula the formula $Bird(tweety)$ is a partial model of the formula simply consisting of one "first-order tuple" $\langle\langle bird, tw; 1\rangle\rangle$. Given a theory I, then correspondingly we look at all the minimal partial models of the formulas in I. We call this the minimal *information* conveyed by the theory I. This is really the approach taken by circumscription as well, since circumscription would rule out the possibility that Tweety is a penguin by minimizing the class of abnormal birds. (A formal comparison between our method and circumscription is made difficult by the fact that we work with partial models and circumscription works with total ones.)

To sum up this discussion, we might say that minimizing knowledge is done by maximizing models. For us, minimizing information is done by minimizing models. Then we use default systems to recreate belief spaces from minimal information.

The organization of the paper is as follows. We motivate our semantics by means of examples taken from Lifschitz, Baader, and Hollunder. We then introduce the formal definition of default models, and state our representation theorems for default systems on Scott domains. Then we present the syntax and semantics of our logic. Next we prove our results on preferential consequences, simultaneously with what we hope are interesting counterexamples to the general laws. Finally we provide a comparison with other approaches in the literature.

There is more to the story of default models. In [24] we show that default models and their logics solve some of the semantic problems with Reiter's default logic. This includes, in particular, a treatment of subjective degrees of belief, as in the work of Bacchus et al. [3]. Further, in [23] we have shown that a propositional modal version of our logic, using weak rather than strong Kleene negation, has a sound and complete equational axiomatization, and an NP-complete validity decision problem.

2. On open defaults

Many of the interesting applications of default logic involve the notion of an *open default*: a default rule of the form

$$\frac{\alpha : \beta_1, \dots, \beta_n}{\gamma}, \tag{1}$$

where some of the formulas may contain free variables. The standard treatment of these defaults, from Reiter's work, is to assume that the default (1) stands for the set of all its ground instances.

Lifschitz, in [17], points out that this treatment can make the effect of a default weaker than intended. He considers the default theory with one axiom $P(a)$ and one default

$$\frac{: \neg P(x)}{\neg P(x)}. \tag{2}$$

This default is intended to express that $P(x)$ is assumed to be false wherever possible. It should allow us to prove the formula

$$\forall x\big(P(x) \leftrightarrow (x = a)\big). \tag{3}$$

However, all that (2) sanctions are the literals $\neg P(t)$ for ground terms t different from a.

Lifschitz notes that the problems of ground terms should be overcome by a treatment of default logic involving the realm of *models*. In that realm, one can talk about domain elements directly, without having to refer to their syntactic names. He cites a result of Guerreiro and Casanova [10] characterizing extensions of closed theories using a model-theoretic fixpoint construction, and then generalizes this construction to the case of open defaults.

Our approach to these problems is similar in spirit to that of Lifschitz, but much more direct. We regard a default rule like (2) not as a syntactic rule at all, but as an "algorithm" for building partial models directly. To see what we mean, assume that P is an atomic predicate. Let M be a set to be used as the universe of a model. We will build models out of "tuples" or *infons* of the form

$$\langle\!\langle P, m; i \rangle\!\rangle, \tag{4}$$

where P is the given predicate symbol, $m \in M$, and i, the *polarity* of the infon, is either 0 or 1, standing for "definite truth" or "definite falsehood". In the given example, we introduce a whole collection of defaults of the form

$$\frac{:\langle\!\langle P, m; 0 \rangle\!\rangle}{\langle\!\langle P, m; 0 \rangle\!\rangle}, \tag{5}$$

where $m \in M$. Our model-building procedure is as follows. We start with a *minimal* model for $P(a)$, namely

$$\big\{\langle\!\langle P, m_a; 1 \rangle\!\rangle\big\}, \tag{6}$$

where m_a is an element of M interpreting the constant a. The defaults (5) can fire to add infons to this minimal model, providing they do not conflict in polarity with $\langle\!\langle P, m_a; 1 \rangle\!\rangle$. Clearly, then, all of the infons $\langle\!\langle P, m; 0 \rangle\!\rangle$ can be added to the model, for $m \neq m_a$. In our theory, the resulting model is the unique default model-theoretic extension of (6). In this extension, the formula (3) holds.

Another, more subtle, problem with Reiter's treatment of open defaults involves the actual reduction of open to closed defaults, which requires a preliminary step of Skolemizing all the axioms and the consequents of all defaults. An example due to Baader and Hollunder makes this point clear, in the context of trying to add defaults to terminological logics. (These are a class of special logics appropriate for knowledge representation.)

Consider a "concept description", or definition, expressing that an adult man is married to a woman, or is a bachelor: In ordinary first-order logic, such a concept description might be

$$AM(x) =_{def} \exists y\big(Spouse(x,y) \land Woman(y)\big) \lor Bachelor(x), \tag{7}$$

where we understand as a background condition that spouses are unique. In knowledge bases, one also has specific facts about specific individuals: Tom, an adult man, is married to a woman named Mary. Instantiating the concept description and representing the specific facts, we get the formula

$$AM(Tom) \land Spouse(Tom, Mary) \land Woman(Mary). \tag{8}$$

For defaults, one assumes, chauvinistically, that if it is consistent that an individual is not a woman, then that individual is not a woman:

$$\frac{:\neg Woman(x)}{\neg Woman(x)}. \tag{9}$$

Baader and Hollunder show that Skolemizing the above formulas using Reiter's method yields anomalous results: Tom is a bachelor married to Mary. The Skolemized version of $AM(Tom)$ is

$$\big(Spouse(Tom, Gordy) \land Woman(Gordy)\big) \lor Bachelor(Tom), \tag{10}$$

where $Gordy$ is a Skolem constant introduced to be the spouse of Tom. Because of the disjunction, (10) does not imply $Woman(Gordy)$. The chauvinistic default can fire, yielding $\neg Woman(Gordy)$. Combined with (10) this yields $Bachelor(Tom)$, which is at odds with Tom's being married to Mary.

Our approach to this problem, and solution in this example, involves defining a first-order logic which is sensitive to *constraints*: laws which govern the behavior of partial models, but which are in the background. (This is an idea taken from situation theory; see Barwise's book [4].) In the above case, situations must respect the constraint that "spouse" is a partial function. To do this technically, we need to cover the basic apparatus of default domain theory.

3. Default domain theory

3.1. Default structures

First we review Scott's [27] idea of *information systems*, which can be thought of as general concrete monotonic "rule systems" for building *Scott domains*: consistently complete algebraic cpo's.

Definition 3.1. An information system is a structure $\underline{A} = (A, Con, \vdash)$ where

- A is a countable set of tokens,
- $Con \subseteq$ FiniteSubsets(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. We will use infons as tokens, as in section 2. We could choose for Con any finite set of token which contains no polarity clash, and we could specify $X \vdash a \Longleftrightarrow a \in X$.

We extend the notion of consistency to arbitrary token sets by enforcing the compactness property, i.e., a set is consistent if every finite subset of it is consistent. Overloading notation a little bit, we still write $y \in Con$, even for infinite y.

Definition 3.2. An (ideal) element of an information system $\underline{A}$ is a set x of tokens which is

1. consistent: $x \in Con$,
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 the above example are exactly *all* of the sets of infons containing no token clashes. This example is not terribly interesting; to get more interesting examples in first-order logic, we introduce *constraints* in section 4.

It is worth noting at this point that *every* consistently complete ω-algebraic partial order can be isomorphically represented as the set of ideal elements of an information system, ordered by subset inclusion. This is Scott's fundamental representation theorem [27], and it accounts for our claim of generality for default information systems. More of this below.

Now we come to the main definitions used in this paper. We introduce the theory of default structures by simply adding a default component to information systems.

Definition 3.3. A default information structure is a tuple

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

where $(A, Con, \vdash)$ is an information system, Δ is a set of triples of members of Con. Each triple (X, Y, Z) is written as $\frac{X:Y}{Z}$.

Notice that the rules in Δ allow us to add a whole “chunk” Z of tokens by default.

The notion of deductive closure associated with information systems plays an important role: the deductive closure $\overline{G}$ of a set G of tokens is the smallest set containing G and closed under $\vdash$. When G is a complicated set we sometimes write $\overline{}[G]$ for $\overline{G}$.

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. There are actually two equivalent definitions.

Definition 3.4 (implicit definition of extensions). Let $\underline{A} = (A, Con, \Delta, \vdash)$ be a default information structure, and x a member of $|\underline{A}|$. For any subset S, define $\Gamma(x, S)$ to be the smallest set $t \subseteq A$ such that

- t is an ideal element;
- $x \subseteq t$;
- if $\frac{X:Y}{Z} \in \Delta$, and $X \subseteq t$ and $S \cup Y$ is consistent, then $Z \subseteq t$.

A set y is said to be an implicit extension of x if $\Gamma(x, y) = y$.

Definition 3.5 (explicit definition of extensions). Let $\underline{A} = (A, Con, \Delta, \vdash)$ 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

$$\begin{aligned}
&\phi(x, S, 0) = x,\\
&\phi(x, S, i+1) = \overline{\phi(x, S, i)} \cup \\
&\qquad\qquad \bigcup \Big\{ Z \mid \frac{X:Y}{Z} \in \Delta \ \&\ X \subseteq \phi(x, S, i) \ \&\ Y \cup S \in Con \Big\},
\end{aligned}$$

y is an explicit extension of x if $\Phi(x, y) = y$.

By Reiter’s proofs, we know that explicit and implicit extensions are the same things. (We have also shown this in [25], where we strengthen the result. See Theorem 3.4 below.) If y is an extension of x, we write $x \delta_{\underline{A}} y$, with the subscript omitted from time to time.

In practice, one often focuses on *normal* systems, where the default rules have the form $\frac{X:Y}{Y}$. This is because in the general case, extensions do not exist, and the use of general defaults is sometimes problematic in other ways. We also consider the so-called *precondition-free* normal defaults, where the set X is empty. We refer to these as PC defaults.

Examples. In Lifschitz' example, let $x = \{\langle\!\langle P, m_a; 1\rangle\!\rangle\}$, and let the defaults be given by Eq. (5) in section 2, regarding each infon $\{\langle\!\langle P, m; 0\rangle\!\rangle\}$ as a singleton set $X_m = \{\langle\!\langle P, m; 0\rangle\!\rangle\}$. Then the unique extension of x is

$$y = x \cup \{\langle\!\langle P, m, 0\rangle\!\rangle : m \neq m_a\}.$$

A more interesting example is the *eight queens* problem. We have in mind in an 8×8 chessboard, so let $8 = \{0, 1, \ldots, 7\}$. Our token set A will be 8×8. A subset X of A will be in Con if it corresponds to an admissible placement of up to 8 queens on the board. For defaults Δ we take the singleton sets

$$\left\{ \frac{: \{\langle i, j\rangle\}}{\{\langle i, j\rangle\}} \mid \langle i, j\rangle \in 8 \times 8 \right\}.$$

We may take $\vdash$ to be trivial: $X \vdash \langle i, j\rangle$ iff $\langle i, j\rangle \in X$. Now, if x is an admissible placement, then the extensions y of x are those admissible placements containing x and so that no more queens may be placed without violating the constraints of the problem.

The last example is intended to guide the reader away from the view of defaults as default logic. In the eight queens problem, it seems desirable to have a language for reasoning about differing placements. For example, given a placement x, is there an extension y which uses all eight queens? This corresponds exactly to our philosophy: default systems are used model-theoretically, and logic is used to describe default models.

We now summarize the main results on extensions from [25].

Theorem 3.1.

1. Extensions always exist for normal systems.
2. If $x \,\delta\, y$ then $y \supseteq x$.
3. If Δ consists only of PC defaults, then $x \,\delta\, y$ iff $y = \overline{}[x \cup \bigcup\{X \mid \frac{:X}{X} \in \Delta \;\&\; y \cup X \in Con\}]$.
4. $x \,\delta\, y$ and $y \,\delta\, z$ implies $y = z$.
5. If $x \,\delta\, y$ and $x \subseteq z \subseteq y$ then $z \,\delta\, y$.
6. If $x \,\delta\, y$ and $x \,\delta\, y'$ then either $y = y'$ or $y \cup y' \notin Con$.

3.2. Abstract default systems

This section gives a treatment of default structures from an abstract, order-theoretic perspective. We think that the abstract presentation of Reiter's work is suggestive of the class of structures to which it really applies: partial models represented as domains. This class is useful in modeling all sorts of higher-order objects, including the type-free lambda calculus. The use of higher-order default systems has yet to be investigated, but we think that there could be many interesting questions about such systems.

To make the paper self contained, we recall some basic domain theoretic definitions.

Definition 3.6. A directed subset of a partial order $(D, \sqsubseteq)$ is a non-empty set $X \subseteq D$ such that for every $x, y \in X$, there is a $z \in X$ with $x \sqsubseteq z$ & $y \sqsubseteq z$. A complete partial order (cpo) is a partial order which has a bottom element and least upper bounds of directed sets. A subset $X \subseteq D$ is bounded (or compatible, consistent) if it has an upper bound in D. A compact (or finite) element a of D is one such that whenever $a \sqsubseteq \bigsqcup X$ with X directed, we also have $a \sqsubseteq y$ for some $y \in X$. We write $\kappa(D)$ for the set of compact elements of D, and let a, b, etc. range over compact elements. A cpo is algebraic if each element of it is the directed least upper bound of a set of compact elements. A cpo is ω-algebraic if it is algebraic and the set of compact elements is countable. A Scott domain is an ω-algebraic cpo in which every compatible subset has a least upper bound. By convention, we write $x \uparrow y$ if the set $\{x, y\}$ is bounded.

The basic theorem, due to Scott, relating Scott domains and information systems, is the following.

Theorem 3.2. The collection of ideal elements associated with an information system forms a Scott domain under inclusion. The compact elements are (the closures of) finite sets X in Con. Moreover, every Scott domain is isomorphic to the collection of ideal elements of some information system. Specifically we can take the token set to be the set $\kappa(D)$, the coherence relation to be the compatibility predicate of D, and $X \vdash a$ iff $a \sqsubseteq \bigsqcup X$.

Our objective is to state a similar representation for an abstract notion of "extension" relation on a Scott domain, involving an abstract presentation of a default system. To do this we need to adjoin an element $\top$ to an arbitrary Scott domain D; we denote this by $D^\top$. The top element $\top$ denotes *inconsistent* information. It is well known that adjoining this element makes $D^\top$ a complete lattice.

Our presentation is given as follows:

Definition 3.7 (abstract defaults). Let $(D, \sqsubseteq)$ be a Scott domain.

1. A *default set* in D is a subset Λ of $\kappa(D)^3$. We call a triple $(a, b, c) \in \Lambda$ a default, and think of it as a rule $\frac{a:b}{c}$, though this is an abuse of notation.
2. Let Λ be a default set, and let u, $v \in D^\top$. The set of Λ-*default consequences* of u with guards in v is

$$\Big\{c \mid (\exists a, b)\big((a, b, c) \in \Lambda \;\&\; a \sqsubseteq u \;\&\; b \uparrow v\big)\Big\}.$$

We denote by $DC_\Lambda(u, v)$ the least upper bound, in $D^\top$, of this set, but we drop the subscript Λ when it is clear from context.

We now present a domain-theoretic version of Reiter's definition of extensions.

Definition 3.8 (abstract extension operator). Let $(D, \sqsubseteq)$ be a Scott domain, and let Λ be an abstract default system over D. Let x, v range over D. Set

$$\Gamma(x, v) = \bigsqcap \bigl\{ z \in D^{\top} \mid x \sqcup DC(z, v) \sqsubseteq z \bigr\}.$$

We say $y \in D$ is an abstract extension of x iff $\Gamma(x, y) = y$.

It is easy to see that $\Gamma(x, v)$ is the least element z of $D^{\top}$ such that (i) $x \sqsubseteq z$, and (ii):

$$\bigsqcup \bigl\{ c \mid (\exists a, b)(a, b, c) \in \Lambda \;\&\; a \sqsubseteq z \;\&\; b \uparrow v \bigr\} \sqsubseteq z.$$

This is almost word for word Reiter's definition of his extension operator. It is also easy to check the following result.

Theorem 3.3. Every default information system determines an extension relation isomorphic to the abstract extension relation on the Scott domain corresponding to the underlying information system, via the correspondence sending a default $\frac{X:Y}{Z}$ to the triple (x, y, z) of compact elements determined by (X, Y, Z); and conversely via the "inverse" correspondence from Scott domains to information systems.

With this result in mind, we make the following observations.

Let u, v, $x \in D^{\top}$, and fix Λ throughout. Consider the operator

$$\Theta(x, u, v) = x \sqcup DC(u, v).$$

Notice that for fixed x and v, that $\Theta(x, u, v)$ is monotonic in u. Then notice, for fixed x and v, the quantity $\Gamma(x, v)$ is the *least prefix-point*, in u, of the operator $\Theta(x, u, v)$ on the complete lattice $D^{\top}$. By the familar Knaster–Tarski theorem, it is also the least fixed point. Moreover, the operator $\Theta(x, u, v)$ is *ω-continuous* in u; that is,

$$\Theta\Bigl(x, \bigsqcup_i u_i, v\Bigr) = \bigsqcup_i \Theta(x, u_i, v)$$

for all linearly ordered sequences u_i, with $i \in \omega$. Now the least fixed point of an ω-continuous operator F on a complete lattice is given by the formula

$$\bigsqcup_i F^i(\bot),$$

where F^i is the i-fold composition of F with itself.

Definition 3.9 (abstract explicit extension operator). Let D be as above, and let $x, v \in D$. Define the operators Π_i as follows:

- $\Pi_0(x,v) = x$;
- $\Pi_{i+1}(x,v) = x \sqcup DC(\Pi_i(x,v),v)$.

Define $\Pi(x,v) = \bigsqcup_i \Pi_i(x,v)$.

By our observation above, $\Pi(x,v)$ is nothing more or less than the least fixpoint of $\lambda u\Theta(x,u,v)$. Therefore,

Theorem 3.4 (coincidence of implicit and explicit extension operators).

$$\Pi = \Gamma.$$

We remarked above that the original result of Reiter was that these two operators determine the same class of extensions, in the special case of default logic, where the Scott domain in question is the set of all deductively closed and consistent sets of sentences of first order logic. The above result – that the operators are identical – holds in all Scott domains.

A consequence of all of the above results is that Theorem 3.1 holds "in the abstract", by the order-isomorphism going from default structures to domains with systems of defaults given abstractly. That is, general facts about extension relations proved concretely using information systems, continue to hold in general abstract domains.

It is natural to wonder if the equation $\Gamma(x,y) = y$ can be simplified. According to the above, $\Gamma(x,v)$ is the least fixed point, in u, of the operator $\Theta(x,u,v)$. Writing this out,

$$\Gamma(x,v) = \bigsqcap\{u \mid x \sqcup DC(u,v) = u\}.$$

This gives that, for $y \neq \top$, that y is an extension of x if and only if

$$y = \bigsqcap\Big\{u \mid u = x \sqcup \bigsqcup\{c \mid (a,b,c) \in \Lambda \;\&\; a \sqsubseteq u \;\&\; b \uparrow y\}\Big\}.$$

This seems to be the simplest equation that we have in general, even for the normal default case. However, in the precondition-free case, simpler formulas do obtain. Suppose that Λ is a PC default system; that is, every default is of the form $(\bot,a,a)$, where $\bot$ is the least element of D. Then Λ can be considered just to be a subset of $\kappa(D)$.

Definition 3.10. The *PC preferential cover* δ_Λ determined by Λ is the binary relation on D given by the condition

$$x\,\delta_\Lambda\,y \Longleftrightarrow y = x \sqcup \bigsqcup\{b \mid (\exists b \in \Lambda)(b \uparrow y)\}.$$

Using this definition, the following result can be proved directly, using the general definitions of domain theory. (By our remarks above, everything but item (3) follows from Theorem 3.1.)

Theorem 3.5. Let D be a Scott domain and Λ be a PC "abstract" default system. We have

1. $\forall x \in D \exists y \in D x \, \delta_\Lambda \, y$.
2. If $x \, \delta_\Lambda \, y$ then $y \sqsupseteq x$.
3. $x \, \delta_\Lambda \, y$ if and only if there is a B, a maximal subset of Λ such that $B \cup \{x\}$ is bounded, which satisfies the equation
$$y = x \sqcup \bigsqcup \{\lambda \mid \lambda \in B\}.$$
4. $x \, \delta_\Lambda \, y$ and $y \, \delta_\Lambda \, z$ implies $y = z$.
5. If $x \, \delta_\Lambda \, y$ and $x \sqsubseteq z \sqsubseteq y$ then $z \, \delta_\Lambda \, y$.
6. If $x \, \delta_\Lambda \, y$ and $x \, \delta_\Lambda \, y'$ then either $y = y'$ or $y \not\uparrow y'$.

To show how these results work, we prove (3) and (5). We do not really need to prove (5), but the argument is fairly elegant in the abstract.

Proof of (3). Suppose $x \, \delta_\Lambda \, y$. Then
$$y = x \sqcup \bigsqcup \{\lambda \in \Lambda \mid \lambda \uparrow y\}.$$
Let B be the set
$$\{\lambda \in \Lambda \mid \lambda \uparrow y\}.$$
We need only check that B is a maximal subset of Λ with $B \cup \{x\}$ bounded. Let $B \subseteq B'$ and suppose $B' \cup \{x\}$ is bounded. Then
$$y = x \sqcup \bigsqcup \{\lambda \mid \lambda \in B\} \sqsubseteq x \sqcup \bigsqcup \{\lambda \mid \lambda \in B'\}.$$
So for every $\lambda \in B'$, we have $\lambda \uparrow y$. Therefore $B' \subseteq B$. This proves one direction; the converse is similar. □

Proof of (5). Let $x \, \delta_\Lambda \, y$ and $x \sqsubseteq z \sqsubseteq y$. Then
$$y = x \sqcup \bigsqcup \{\lambda \in \Lambda \mid \lambda \uparrow y\}.$$
Since $x \sqsubseteq z$, we have
$$y \sqsubseteq z \sqcup \bigsqcup \{\lambda \mid \lambda \uparrow y\} \sqsubseteq y \sqcup \bigsqcup \{\lambda \mid \lambda \uparrow y\} = y.$$
Therefore $z \, \delta_\Lambda \, y$ by definition. □

A direct application of the above definition is to the problem of defaults in lexical semantics of natural language, found in Young's paper [30] on nonmonotonic sorts. There D is chosen to be a lexical hierarchy, say for verbs, and Λ is "linguist-chosen" to reflect, for example, preferred default verb endings.

The analogy between Theorems 3.5 and 3.1 suggests that the extension relation of a PC default information structure is a concrete representation of any PC preferential covering relation, and this is indeed the case. Moreover, the proof is easy given our characterizations above.

Theorem 3.6. If Λ consists of PC defaults, then y is an extension of x if and only if $x \, \delta_\Lambda \, y$.

Proof. Consider $\Theta(x,u,v)$ for precondition-free defaults. We have

$$\Theta(x,u,v) = x \sqcup DC(u,v) = x \sqcup \bigsqcup\{b \mid b \in \Lambda \;\&\; b \uparrow v\}.$$

This means that Θ does not depend on u and so

$$\Gamma(x,v) = x \sqcup \bigsqcup\{b \mid b \in \Lambda \;\&\; b \uparrow v\},$$

from which the desired conclusion follows. □

Remarks.

- We will use the form of Theorem 3.5 to help us prove one of the Cautious Monotony laws for precondition-free systems.
- The characterization in Theorem 3.6 does not generalize to normal systems. That is, we could define

$$x \, \delta_\Lambda \, y \iff y = x \sqcup DC(y,y).$$

  If y is an extension of x, then it follows that $x \, \delta_\Lambda \, y$, but not conversely, even if Λ is normal. To see this, let D be a Scott domain consisting of 3 elements $\{\bot, 0, 1\}$, with 0 and 1 incompatible. Let $\Lambda = \{(0,0,0)\}$ (intuitively the default $\frac{0:0}{0}$). Then $0 = \bot \sqcup DC(0,0)$, but 0 is not an extension of $\bot$. Further, neither maximal nor minimal solutions to the above equation work either. For a counterexample to the minimal version, add the precondition-free default $\frac{\bot:1}{1}$ to the above system. Then 0 is a minimal solution, but is still not an extension.

4. Default models for first-order logic

Having covered the basics of default domain theory, we introduce our logic for default reasoning. To begin, let V be a countable set of variables, and fix a first-order relational signature. (The restriction to relational signature is just for convenience.) The syntax of our language $\mathcal{L}$ is just that of ordinary FOL. We thus have atomic formulas of the form $\sigma(p_1,\ldots,p_n)$, where σ is an n-ary relation symbol, and the p_i may be constants or variables.

The language $\mathcal{L}$ is generated by the grammar

$$\varphi ::= \textbf{true} \mid \textbf{false} \mid \sigma(p_1,\ldots,p_n) \mid x = y \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \neg\varphi \mid \exists x\varphi \mid \forall x\varphi.$$

Remarks on the language:

1. The operator $\neg$ stands for strong Kleene negation. If we have $\neg\varphi$ holding in some situation, then we can be sure φ will never be realized in a larger situation.
2. We introduce two other formulas by abbreviation: $(\varphi \to \psi) = \neg\varphi \vee \psi$, and the biconditional $\leftrightarrow$ similarly.
3. We could add a modality B to $\mathcal{L}$. $\mathsf{B}\varphi$ is read "φ is believed". The meaning of $\mathsf{B}\varphi$ is that with respect to a situation w, φ holds in all extensions of w. This was done in [23].

The semantics of this language is its nonstandard feature. Our structures will be default information systems based on some particular set of individuals M, but we first have to assume some *background constraints* on any relations which are going to be holding in such sets M. These constraints will be used to generate the monotonic entailment relation $\vdash$ in the default structure. (The defaults themselves can be arbitrary, as long as they are precondition-free.) 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.[3] We can also interpret function symbols to be actual functions on M, but this will not be needed in what follows.

Now, tokens of our information system will be infons of the form

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

where $m_j \in M$, 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

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

where the R_j are relations on M of the appropriate arities, such that $\mathcal{M}$ is a model of C, and such that if $\langle\langle \sigma_j, m_1, \ldots, m_n; 1 \rangle\rangle \in s$, then the corresponding $R_j(m_1, \ldots, m_n)$ is true. Similarly, if $\langle\langle \sigma_j, m_1, \ldots, m_n; 0 \rangle\rangle \in s$, then the corresponding $R_j(m_1, \ldots, m_n)$ is false.

We have a choice in the treatment of the equality relation. One possibility is to assume that it is interpreted as a *partial congruence relation* on M. In that case, we would use the special symbol E for this relation. To get it to be a partial congruence, we use special constraints: first order sentences stating that E is an equivalence relation, and that E respects relations in the usual way. For the examples in this

[3] In terms of philosophy of language, we are taking constants to be rigid designators; in terms of theorem proving, they are "non-flexible".

paper, this distinction between true identity and congruential identity does not matter. We have therefore chosen to simplify the presentation by assuming that equality is always interpreted as the identity.

An admissible set of infons is *total* if it is maximal in the subset ordering on sets of infons. Intuitively, this is an acceptable truth assignment, or possible world, in the 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.

Definition 4.1. 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 **true**. 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 $C =$ **true**, 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.

A more interesting constraint comes from considering the married bachelor problem in section 2. Take C to be the single constraint

$$\forall xyz\big(Spouse(x,y) \wedge Spouse(x,z) \rightarrow y = z\big).$$

This constraint will force partial models, or situations, to respect the unique spouse property.

It can be checked that our definition of $\vdash$ satisfies Scott's conditions for a monotonic entailment relation. (To be absolutely correct, we should only consider those tokens σ such that $\{\sigma\}$ is admissible. But this requirement of Scott systems is not important for our purposes.)

We are finally ready to give a interpretation to our logic. Select a default structure $\underline{A}(M,C)$ as above. Let s be a situation of $\underline{A}(M,C)$, and let $\alpha : V \rightarrow M$. Let x and y range over situations in $\underline{A}$. We define two relations $\models$ and $=\!|$ between $|A|$ and $\mathcal{L}$ by simultaneous induction:

- $(s, \alpha) \models$ **true** always;
- $(s, \alpha) \dashv$ **false** always;
- $(s, \alpha) \models \sigma(v_1, \ldots, v_n)$ iff $\langle\langle \sigma, \alpha(v_1), \ldots, \alpha(v_n); 1 \rangle\rangle \in s$; (If some of the v_i are constraints, we use instead of $\alpha(v_i)$ the fixed interpretation of that constant.)
- $(s, \alpha) \dashv \sigma(v_1, \ldots, v_n)$ iff $\langle\langle \sigma, \alpha(v_1), \ldots, \alpha(v_n); 0 \rangle\rangle \in s$;
- $(s, \alpha) \models x = y$ iff $\alpha(x) = \alpha(y)$;
- $(s, \alpha) \dashv x = y$ iff $\alpha(x) \neq \alpha(y)$;
- $(s, \alpha) \models \varphi \vee \psi$ iff $(s, \alpha) \models \varphi$ or $(s, \alpha) \models \psi$;
- $(s, \alpha) \dashv \varphi \vee \psi$ iff $(s, \alpha) \dashv \varphi$ and $(s, \alpha) \dashv \psi$;
- $(s, \alpha) \models \varphi \wedge \psi$ iff $(s, \alpha) \models \varphi$ and $(s, \alpha) \models \psi$;
- $(s, \alpha) \dashv \varphi \wedge \psi$ iff $(s, \alpha) \dashv \varphi$ or $(s, \alpha) \dashv \psi$;
- $(s, \alpha) \models \neg\varphi$ iff $(s, \alpha) \dashv \varphi$;
- $(s, \alpha) \dashv \neg\varphi$ iff $(s, \alpha) \models \varphi$;
- $(s, \alpha) \models \exists v\varphi$ iff for some $m \in M$, $(s, \alpha[v \leftarrow m]) \models \varphi$;
- $(s, \alpha) \dashv \exists v\varphi$ iff for all $m \in M$, $(s, \alpha[v \leftarrow m]) \dashv \varphi$;
- $(s, \alpha) \models \forall v\varphi$ iff for all $m \in M$, $(s, \alpha[v \leftarrow m]) \models \varphi$;
- $(s, \alpha) \dashv \forall v\varphi$ iff for some $m \in M$, $(s, \alpha[v \leftarrow m]) \dashv \varphi$.

The relations $\models$ and $\dashv$ are read "positively satisfies" and "negatively satisfies", respectively. As usual, we write $A(M, C) \models \varphi$ iff for all (s, α) we have $(s, \alpha) \models \varphi$.

Our semantics is a standard one for Kleene 3-valued logic; see [5] for a complete treatment. In particular, the default play no role; the nonmonotonic effect of the defaults will be considered in the next section. We need the following result about the semantics.

Lemma 4.1 (persistence). For a sentence φ of $\mathcal{L}$, if $s \models \varphi$ and $s \subseteq t$, then $t \models \varphi$.

The proof is straightforward, using a double induction over the structure of formulas.

Another useful result, whose proof is standard, is that

Lemma 4.2. Every formula is logically equivalent to one in which negation appears only on atoms.

(By logical equivalence, we mean that two formulas are both positively and negatively satisfied the same way in all structures.)

We will be concerned in the sequel with the existence of minimal models for certain formulas of our logic. Domain theory helps us to find a sublogic of first-order logic for which we can always guarantee that a satisfiable formula has a minimal model. Every domain (in fact algebraic cpo) has a "natural" positive logic associated with it. This logic is given by the *Scott topology* of the domain.

Definition 4.2. Let $(D, \sqsubseteq)$ be a cpo. A set $U \subseteq D$ is said to be *Scott open* if for any directed $X \subset D$, we have $\bigsqcup X \in U$ iff $U \cap D \neq \emptyset$.

One checks readily that under this definition of "open", that the collection of open subsets of a cpo form a topological space. Such a space must contain $\emptyset$ and D, and be closed under finite intersections and arbitrary unions. Furthermore, we can regard open sets as being "properties" of domain elements. The definition says that if an element has a certain property, then we can discover that the property holds by testing a "sequence" of finite elements which "converges" to the given element. (In general, sequence really means directed set, and "converges to an element" means that the element is the least upper bound of the set.) After a finite time, we find that the element does indeed have the property. Such properties are sometimes called "affirmable" [29].

It is straightforward to prove the following in any algebraic cpo D.

Theorem 4.1 (compactness in the Scott topology).

1. For each finite element $f \in D$, the principal compact filter $\uparrow f = \{u \mid f \sqsubseteq u\}$ is open.
2. Every open set U is the union of the principal compact filters generated by the compact elements of U.
3. Every compact open set X is a finite union of such compact principal filters. (*Compact here means the topological usage: every covering of X by open sets has a finite subcovering.*)

These results should guide us to a positive sublogic of FOL for which each closed formula defines an open set in the Scott topology. In fact, we can do the following:

Definition 4.3. The *positive* fragment of FOL is the least set of formulas including positive and negated relational atoms (not including equations or inequations), and closed under conjunction, negation, and existential quantification.

Lemma 4.3. Every (closed) formula in the positive fragment of FOL defines a Scott open set in the topology of the domain defined by a given information structure. More precisely, for fixed φ, α, the set $\{s : (s, \alpha) \models \varphi\}$ is open.

Proof (*sketch*). One first notices that basic literals define open sets, and then uses induction on the structure of the formula. We do not need to consider the case of negation or universal quantification, so we do not use the $=\!\!|$ relation. The inductive step for existential quantification uses the fact that the open sets in any topology are closed under arbitrary unions. $\square$

Notice that in general, not every open set can be written as the set defined by some positive formula. This is because (at least for a countable vocabulary and structure) there are uncountably many open sets, but only countably many formulas.

5. On nonmonotonic consequence relations

5.1. Definition and examples

We turn to our notion of "nonmonotonic consequence", and begin with definitions.

Definition 5.1. Let φ be a sentence in positive first-order logic. Let $\underline{A} = A(M, C)$ be a default information system. A *minimal model* of φ is an ideal element m of $\underline{A}$ such that

- $m \models \varphi$;
- m is inclusion-minimal in the set of n with $n \models \varphi$.

Definition 5.2. Let φ be a positive sentence, and β be an arbitrary sentence in first-order logic. Let $\underline{A} = A(M, C)$ be a default information system as above.

We say that $\varphi \mathrel{|\!\!\sim}_{\underline{A}} \beta$ if for all minimal situations s such that $s \models \varphi$, we have that

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

We will see the reasons for these definitions when we study the general properties of nonmonotonic consequence relations below. We should explain at this point, though, one reason for requiring formulas on the left of the $\mathrel{|\!\!\sim}$ sign to be positive. This comes from the Scott topology, and the fact that open sets define "affirmable" properties. We think of the formula φ as defining a property for which we can have evidence. Such evidence might be observations of birds, for example. It does not make much sense for "evidence" to be given in the form of universal statements. On the other hand, leaping to a conclusion β quite often is an inductive leap: from observations of birds, we conclude that they all fly. So we do not in general require β to be positive. Incidentally, the same restriction to positive formulas guarantees the existence of minimal models for satisfiable formulas, a technical condition needed for our results, and one which does not obtain for partial models of FOL in general.

5.2. Open defaults revisited

Here now are case studies showing that our idea of nonmonotonic consequence is reasonable. We use our notion of nonmonotonic consequence to resolve Lifschitz' problem, and the problem of Baader and Hollunder from section 2.

Example 1. Consider Lifschitz' example from section 2. Let $\underline{A} = \underline{A}(M, C)$ be a default structure as above, with C being the formula **t**. We translate the default $\frac{:\neg P(x)}{\neg P(x)}$ as the collection of defaults of the form

$$\frac{:\{\langle\!\langle P, m; 0\rangle\!\rangle\}}{\{\langle\!\langle P, m; 0\rangle\!\rangle\}},$$

where m can be an arbitrary element of M. We then have

$$P(a) \mathrel{|\!\!\sim}_{\underline{A}} \forall x(P(x) \leftrightarrow x = a).$$

Notice how $P(a)$ is a typical bit of evidence.

Example 2. Consider the case of the married bachelor. We are going to treat this problem model-theoretically, to see if we can rule out the anomaly. (Refer to section 2 for the anomaly caused by the Skolem technique.)

First of all, notice that Skolemization cannot make use of the background information

$$C = \forall xyz\big(Spouse(x, y) \wedge Spouse(x, z) \rightarrow y = z\big)$$

that spouses are unique. (This is true even if we add Skolem axioms for the introduced constant as part of the base theory.) Our system *can* make use of the uniqueness constraint by virtue of the model theory. Let $\underline{A}$ be a default information structure for the above language, using constraint C, and with set M of individuals. Let ma be the fixed element of M interpreting the constant $Mary$, and tom likewise for Tom. For the defaults we simply use

$$\frac{: \langle\!\langle woman, m; 0\rangle\!\rangle}{\langle\!\langle woman, m; 0\rangle\!\rangle}$$

for each m in M. Now suppose that w is a minimal situation satisfying our specific world description

$$\phi = AM(Tom) \wedge Spouse(Tom, Mary) \wedge Woman(Mary).$$

Then in particular w satisfies $AM(Tom)$, so either (i) $\langle\!\langle bachelor, tom; 1\rangle\!\rangle \in w$ or (ii) for some $m \in M$, $\langle\!\langle spouse, tom, m; 1\rangle\!\rangle \in w$, and $\langle\!\langle woman, m; 1\rangle\!\rangle \in w$. In case (ii), we know that w must also respect the background constraint, so m is the only element of M with the property of being a spouse of tom. Since w also satisfies $spouse(Tom, Mary)$ we must have $m = ma$. It is easy to see that w will have a unique extension x; this is because no defaults themselves conflict. Since $\langle\!\langle woman, ma; 1\rangle\!\rangle \in w$, the extension x cannot contain the information that Mary is not a woman. However, it will contain the tokens stating that all the other elements of M are non-women, since w is minimal.

Now suppose that (ii) is not the case. Then for any $m \in M$, either $\langle\!\langle woman, m; 1\rangle\!\rangle$ is not in w, or $\langle\!\langle spouse, tom, m; 1\rangle\!\rangle$ is not in w. What about the individual ma? Since $\langle\!\langle woman, ma; 1\rangle\!\rangle \in w$, the first possibility is ruled out, and since $\langle\!\langle spouse, tom, m; 1\rangle\!\rangle \in w$, so is the second. This is a contradiction, and (ii) must be the case. The situation in the extension is thus that Mary is a woman; everyone else is a non-woman, and Tom is married to Mary. We cannot infer anything about Tom's being a bachelor, since we have no constraint about bachelors having no spouses (to see this, replace "bachelor" with "gardener").

In other words, in $\underline{A}$ we have

$$\phi \mathrel{|\!\sim}_{\underline{A}} \Big(\forall x\big(x \neq Mary \to \neg Woman(x)\big)\Big) \wedge Spouse(Tom, Mary) \wedge Woman(Mary).$$

Once again, the formula ϕ is an affirmable one.

A general procedure for building a default information system to model open defaults is the following. Consider the constraints C and the open default

$$\frac{\alpha(x_1,\ldots,x_n) : \beta(x_1,\ldots,x_n)}{\gamma(x_1,\ldots,x_n)},$$

where the only free variables in any of the formulas are among the x_i. We assume the three formulas are conjunctive; that is, built from atoms or negated atoms using conjunction alone. (If we are working over finite structures, universal quantification would also be allowed.) The reason for this assumption is when we build extensions, we always add conjunctive information to our models by adjoining finite sets of infons. In other words, we regard the syntactic form of open defaults as a schema for specifying which finite chunks of information will be added to our models.

Select a set M and a structure $\underline{A}(M, C)$. Use for defaults the collection

$$\frac{m(\alpha) : m(\beta)}{m(\gamma)},$$

where each $m(\varphi)$ is a minimal model of $\exists x\varphi(x)$, but also such that for some assignment τ to the free variables of the defaults, $(m(\alpha), \tau) \models \alpha$, $(m(\beta), \tau) \models \beta$, and $(m(\gamma), \tau) \models \gamma$. The idea here is that we really want to consider sets of "ground" instances of infons. We can have not only ground instances where constants are named, but also unnamed elements of a structure.

5.3. *Preferential consequence*

Our objective in this section is to consider the basic Kraus–Lehmann–Magidor axioms for any "reasonable" notion of preferential entailment. A "core" set of laws is the following:

1. (**Reflexivity**) $\varphi \mathrel{|\!\sim} \varphi$.
2. (**Left Logical Equivalence**) If $\models (\varphi \leftrightarrow \psi)$, and $\varphi \mathrel{|\!\sim} \beta$, then $\psi \mathrel{|\!\sim} \beta$.
3. (**Right Weakening**) If $\models (\alpha \to \beta)$ and $\varphi \mathrel{|\!\sim} \alpha$, then $\varphi \mathrel{|\!\sim} \beta$.
4. (**Cut**) If $\varphi \wedge \alpha \mathrel{|\!\sim} \beta$, and $\varphi \mathrel{|\!\sim} \alpha$, then $\varphi \mathrel{|\!\sim} \beta$.
5. (**Cautious Monotony**) If $\varphi \mathrel{|\!\sim} \alpha$ and $\varphi \mathrel{|\!\sim} \beta$, then $\varphi \wedge \alpha \mathrel{|\!\sim} \beta$.
6. (**Or**) If $\alpha \mathrel{|\!\sim} \gamma$ and $\beta \mathrel{|\!\sim} \gamma$, then $\alpha \vee \beta \mathrel{|\!\sim} \gamma$.

Our interpretation of the laws is as follows. Consider a law like (2). We interpret it to mean: For all structures $\underline{A}$, if $\underline{A} \models (\varphi \leftrightarrow \psi)$, and $\varphi \mathrel{|\!\sim}_{\underline{A}} \beta$, then $\psi \mathrel{|\!\sim}_{\underline{A}} \beta$.

The laws 1, 2, 3 follow easily from the definition of $\mathrel{|\!\sim}$. (Note that (1) requires persistence, Lemma 4.1.) We consider first the Cut law (4), and then the Cautious Monotony law (5). We treat each law in turn. First we show that in general normal structures, each is false. We then consider restrictions on the form of the law, and on semantic structures, which make it true. In each case, we try to give intuitive explanations which in some way show that the counterexamples are in fact reasonable. Lastly we treat the Or law. Although this law is true, there is an objection to our definition of $\mathrel{|\!\sim}$ involving non-exclusive disjunctions. We consider a variant definition of $\mathrel{|\!\sim}$ and a variant of the Or law, called the Cautious Or law, which is verified by the variant definition of $\mathrel{|\!\sim}$.

To increase readability, the counterexamples are stated using propositional letters instead of first-order formulas. Recall that $\neg$ denotes strong negation, on the syntactic side. We use $\overline{f}$ to denote the polarity 0 infon corresponding to the negated atom $\neg f$.

Counterexample to (4) using precondition-free defaults. Let φ be $w \vee \neg f$, $\alpha = w \vee f$, and $\beta = f$. Take for defaults

$$\frac{:f}{f} \quad \text{and} \quad \frac{:w}{w}.$$

The minimal models of φ are $\{w\}$ and $\{\overline{f}\}$. For $\varphi \wedge \alpha$ the unique minimal model is $\{w\}$. This has the unique extension $\{w, f\}$, so $\varphi \wedge \alpha \mathrel{|\!\sim} \beta$. Also, the extensions of the minimal models of φ are $\{w, \overline{f}\}$, respectively. In both of these w holds, so that $\varphi \mathrel{|\!\sim} \alpha$. But φ does not $\mathrel{|\!\sim} \beta$, since we have the extension $\{w, \overline{f}\}$.

The problem here is that by moving to a minimal model of $\varphi \wedge \alpha$ we are forgetting the information in the models $\{w, f\}$ and $\{w, \overline{f}\}$ which we had when we were figuring out the conjunction. The second of these models would block the extension $\{w, f\}$.

Is this counterexample a realistic one? Imagine that w stands for a property that the typical bird has, like "wingspan less than 6 feet"; and f stands for the property that the typical bird has, like "wingspan less than 6 feet"; and f stands for the property of flying. Using a new atom b for "bird", we could reason with the KLM laws as follows: Suppose that birds normally fly, and birds normally have wingspans less than 6 feet. Using intuitive reasoning, it seems that from $b \wedge (w \vee \neg f)$ we could jump to the conclusion $b \wedge w$. It also seems reasonable to accept $b \wedge w \mathrel{|\!\sim} f$. But from $b \wedge (w \vee \neg f) \mathrel{|\!\sim} b \wedge w$ we get by weakening $b \wedge (w \vee \neg f) \mathrel{|\!\sim} (b \wedge w) \vee (b \wedge f)$. Let α be the formula $(b \wedge w) \vee (b \wedge f)$, ϕ be the formula $b \wedge (w \vee \neg f)$, and β be f. Then $\phi \wedge \alpha$ is equivalent to $b \wedge w$, from which we conclude f. But from $(b \wedge w) \vee (b \wedge \neg f)$ it does not seem reasonable to conclude f because of the case $b \wedge \neg f$.

Where is the problem in the intuitive reasoning? First, it could be in the acceptance of $b \wedge w \mid\!\sim f$. By shortening wingspans, we are lessening the probability of flying. But almost all of the small birds we are aware of do fly[4]. Second, the problem could be in the acceptance of the conclusion of having a normal wingspan from the disjunction of that property with that of not flying. From our experience, at least in AI, non-flying birds are rather large. But most of them do not have six-foot wingspans, as do eagles, condors, and vultures. Finally, we could accept flying from the disjunction of non-flying with the property of having a normal wingspan. This is probably what we should do, since the non-flying birds are very rare compared to the population with normal wingspans. But from a purely logical perspective, it seems incorrect.

To explore the situation further, consider removing the disjunction in α.

An example using normal defaults. Let n stand for "nixonian person", q for "quaker", and p for "pacifist". Let φ be $n \vee q$, $\alpha = q$, and $\beta = p$. Take for defaults

$$\frac{n : \overline{p}, q}{\overline{p}, q} \quad \text{and} \quad \frac{q : p}{p}.$$

Assume that $\{p, \overline{p}\} \vdash \emptyset$. (The notation $X \vdash \emptyset$ means that X is not in Con.)

We have $\varphi \mid\!\sim \alpha$, and $\varphi \wedge \alpha \mid\!\sim \beta$, for the same reasons as above. But we do not have $\varphi \mid\!\sim \beta$ because of the minimal model n of φ and its extension $\{n, \overline{p}, q\}$, in which the second default rule is blocked. This shows that we cannot take the cut formula α to be conjunctive. And in general, it seems that we would not want to make the inference, unless it were a statement about probabilities. That is, the number of Nixon-like individuals is very small relative to the number of pacifists and quakers. Once again, the conclusion of pacifism is warranted on probabilistic grounds. But that situation may not arise in conversation. On hearing the following:

> George is either a Nixonian or an ordinary Quaker.

the listener, even knowing the various approximate entailments above, will not jump to the conclusion that George is a pacifist, because the conversational situation overrides the fact that George may have been randomly selected from the Nixonians together with the Quakers. Instead, he or she will assign roughly equal degrees of belief to each of the two disjuncts, because otherwise the statement would not have been made.

We conjecture that when defaults are precondition-free, the cautious cut rule is valid with arbitrary φ and conjunctive α. What about the other possibility: φ is conjunctive, and α is arbitrary? This works in general.

Theorem 5.1. The cautious cut rule holds in the case that φ has a unique minimal model.

[4] An exception could be made for babies.

Proof. Let m be the unique minimal model of φ. (In other words, the compact generator of the open set $\{n \mid m \sqsubseteq n\}$.) Let f be an extension of m. We want to show that $f \models \beta$, assuming that $\varphi \mid\!\sim \alpha$ and $\varphi \wedge \alpha \mid\!\sim \beta$. Let n be a minimal model of $\varphi \wedge \alpha$, such that $n \sqsubseteq f$. This is possible to find, since $f \models \varphi$ by persistence from m, and $f \models \alpha$ by the hypothesis $\varphi \mid\!\sim \alpha$. Then

$$m \sqsubseteq n \sqcup m \sqsubseteq f.$$

We will be done if we can find a minimal model k of $\varphi \wedge \alpha$ such that $m \sqsubseteq k \sqsubseteq n \sqcup m$, because then $m \sqsubseteq k \sqsubseteq f$, so that f is an extension of k, by Theorem 3.1(5). But the set of models of $\varphi \wedge \alpha$ containing $n \sqcup m$ as an element is nonempty and open, and further is a subset of $\uparrow m$. So there is a compact generator k of this set in between m and $m \sqcup n$, as desired. □

In fact, more is true. Say that an open set is *disjoint* if for any two distinct compact generators x and y, we have that x and y are inconsistent. A (positive) formula is disjoint if the set of its positive models is a disjoint open set. (A disjoint open set is called a *stable neighborhood* in domain theory; see [31].)

Theorem 5.2. The cautious cut rule holds when φ is disjoint.

Proof. As before, let n be a minimal model of $\varphi \wedge \alpha$, such that $n \sqsubseteq f$. We have

$$m \sqsubseteq n \sqcup m \sqsubseteq f.$$

We will be once again done if we can find a minimal model k of $\varphi \wedge \alpha$ such that $m \sqsubseteq k \sqsubseteq n \sqcup m$. But now let k be a minimal model of $\varphi \wedge \alpha$ with $k \sqsubseteq n \sqcup m$. Let r be a minimal model of φ with $r \sqsubseteq k$. Then r and m are compatible, so by the disjointness property, $r = m$. Therefore $m \sqsubseteq k \sqsubseteq n \sqcup m$ as before. □

Counterexample to the Cautious Monotony Law (5). Consider the token set $\{p, f, \overline{f}\}$. Intuitively we want p to stand for “penguin”, and f for “fly”. Consider the information system generated by $\{p\} \vdash \overline{f}$ and $\{f, \overline{f}\} \vdash \emptyset$. Choose for a default

$$\frac{:f}{f}.$$

Then we have **true** $\mid\!\sim f$ and **true** $\mid\!\sim f \vee p$, but $f \vee p$ does not $\mid\!\sim f$. This counterexample seems more cogent than the previous ones, even though a probabilistic analysis as above validates the law. In other words, from “birds normally fly”, it follows by weakening that birds normally fly or are penguins. If Cautious Monotony held, then flying should follow from the disjunction of “bird and fly” with “bird and penguin”. This seems like incorrect reasoning by cases.

Next, we show that even though Cautious Monotony does not hold in general, it holds in the PC case when the formula α which is accumulated into the hypothesis is conjunctive.

Theorem 5.3. The Cautious Monotony law holds in the PC case when α has a unique minimal model.

Proof. We work with information systems. Consider a PC default structure $\underline{A}$ and let φ and α define open sets in the Scott topology. Assume that α is conjunctive, that $\varphi \mathrel{|\!\sim} \alpha$, and that $\varphi \mathrel{|\!\sim} \beta$. Let e be any extension of a minimal model $m(\varphi \wedge \alpha)$ of $\varphi \wedge \alpha$. We want to show that $e \models \beta$. Let Λ be the collection of elements $\overline{X}$, where $\frac{:X}{X}$ is a default rule of $\underline{A}$. Then by the definitions and results in section 3,

$$e = m(\varphi \wedge \alpha) \sqcup \bigsqcup \{\lambda \in \Lambda \mid \lambda \uparrow e\}.$$

Let $G(\varphi \wedge \alpha, e)$ denote the set $\{\lambda \in \Lambda \mid \lambda \uparrow e\}$.

We may now decompose $m(\varphi \wedge \alpha) = m(\varphi) \sqcup m(\alpha)$, where $m(\varphi)$ is some minimal model of φ, and where $m(\alpha)$ is the unique minimal model of α. (This is a straightforward exercise in partial orders.)

We next claim that

$$f = m(\varphi) \sqcup \bigsqcup G(\varphi \wedge \alpha, e)$$

is an extension of $m(\varphi)$. To see this, we apply Theorem 3.5(3). Notice that the set $G(\varphi \wedge \alpha, e)$ is compatible with $m(\varphi)$. So the only way that f could not be an extension of $m(\varphi)$ is that there is some maximal $G \subseteq \Lambda$ properly including $G(\varphi \wedge \alpha, e)$, and compatible with $m(\varphi)$, so that

$$g = m(\varphi) \sqcup \bigsqcup G$$

is an extension of $m(\varphi)$. But since $\varphi \mathrel{|\!\sim} \alpha$ by hypothesis, we have $g \models \alpha$, and so $m(\alpha) \sqsubseteq g$, because α has a unique minimal model. We therefore have $m(\varphi) \sqcup m(\alpha) \sqsubseteq g$, so that G is compatible with $m(\varphi) \sqcup m(\alpha) = m(\varphi \wedge \alpha)$. Therefore, $G(\varphi \wedge \alpha, e)$ would not be a maximal set of defaults compatible with $m(\varphi \wedge \alpha)$, and hence e would not have been an extension of $m(\varphi \wedge \alpha)$. Thus, f is an extension of $m(\varphi)$.

The result now follows, as $f \models \beta$ by hypothesis, and $f \sqsubseteq e$, so that $e \models \beta$ by persistence. This completes the proof. $\square$

This result does not extend to normal default structures: the nonmonotonic consequence relation induced by a normal default structure does not in general have the cautious monotony property, even when all formulas are conjunctive. Consider the following normal default structure $(A, \mathit{Con}, \vdash, \Delta)$, where

$$A = \{a, b, c\},$$
$$\Delta = \left\{ \frac{:b}{b}, \frac{b:a}{a}, \frac{a:c}{c} \right\},$$
$$\{a, b, c\} \vdash \emptyset.$$

There is a unique extension for $\emptyset$: $\{a, b\}$. There are, however, two extensions for $\{a\}$:

$$\{a, b\},\ \{a, c\}.$$

We have, therefore, **true** $\mid\!\sim a$, **true** $\mid\!\sim b$, but $a \not\mid\!\sim b$.

Finally, we consider reasoning by cases. Our definitions make the next result trivial.

Theorem 5.4. The Or law

$$\alpha \mid\!\sim \gamma \text{ and } \beta \mid\!\sim \gamma \text{ imply } \alpha \vee \beta \mid\!\sim \gamma$$

holds in general.

There is a possible objection at this point to our definition of $\mid\!\sim$. The conjunction of two formulas α and β is a strengthening of both of them. If we are reasoning by cases, then the "both" part of "either α or β or both" may defeat the conclusions following nonmonotonically from α and also the ones following from β. For example, take a default structure with a trivial $\vdash$ relation over the token set $\{a, b\}$. For defaults, take

$$\frac{a : \overline{b}}{\overline{b}},\ \frac{b : \overline{a}}{\overline{a}}.$$

Then we have $a \mid\!\sim \neg a \vee \neg b$ and $b \mid\!\sim \neg a \vee \neg b$. By the Or law, we have $a \vee b \mid\!\sim \neg a \vee \neg b$. However, consider the least model of $a \wedge b$, which is $\{a, b\}$. None of the defaults are applicable in this situation. So $a \wedge b$ defeats the previous conclusions $\neg a$ and $\neg b$.

This problem suggests that we might want to use a more conservative notion of $\mid\!\sim$.

Definition 5.3. Let φ be a sentence in positive first-order logic. Let $\underline{A} = A(M, C)$ be a default information system. An **almost-minimal model** of φ is an ideal element m of $\underline{A}$ such that

- $m \models \varphi$;
- m is the least upper bound of a finite collection of n's inclusion-minimal in the set $\{n \mid n \models \varphi\}$.

Then we say that (in a structure $\underline{A}$) $\phi \mid\!\sim^a_A \psi$ iff for any almost-minimal model m of ϕ, all default extensions of m satisfy ψ. (The superscript a stands for "almost".)

One can check that the previous results still hold under the $\mid\!\sim^a$ interpretation. Further, it is straightforward to show the following.

Theorem 5.5. The Cautious Or law

$$\alpha \mid\!\sim^a \gamma \quad \text{and} \quad \beta \mid\!\sim^a \gamma \quad \text{and} \quad (\alpha \wedge \beta) \mid\!\sim^a \gamma \quad \text{imply } (\alpha \vee \beta) \mid\!\sim^a \gamma$$

holds in the "almost-minimal" sense $\mid\!\sim^a$.

Notice that default logic itself does not support case analysis. Our theory, even though we use a conservative notion of "cases", does support this notion.

6. A comparison with other approaches

In this section we discuss probabilistic and cumulative-model approaches to the KLM laws, in an attempt to understand why our semantics does not satisfy them.

Probabilistic analyses of defaults, such as those in Adams [1], and subsequently Pearl [21], among many others, have pointed to the laws (1)–(6) (section 5.3) as "core laws" which should hold in any reasonable calculus of approximate reasoning.

Consider the first counterexample to the Cut law. Simple numerical experiments confirm that on a probabilistic basis, acceptance of the conclusion of flying is warranted even when not flying is one of the cases in the premise. Why, then, does this conclusion seem unintuitive? The answer lies, we think, in the use of default reasoning in natural language. As Pearl points out:

> In the logical tradition, defaults are interpreted as conversational conventions, rather than as descriptions of empirical reality ... The purpose ... is not to convey information about the world, but merely to guarantee that, in subsequent conversations, conclusions drawn by the informed will match those intended by the informer [21, section 1].

We agree with this analysis, but we also think that from the use of default conditionals in conversation, listeners may quite possibly create multiple world models in which different models get assigned rough and sometimes inaccurate degrees of belief. This would be the case above, where the possibility of non-flying birds gets assigned more weight than would be warranted empirically. It would also account for the Nixon–Quaker example, as we mentioned above.

Adams' ϵ-semantics verifies the laws, but involves interpreting defaults in a limit sense: $\phi \mid\!\sim \psi$ is interpreted as saying that the conditional probability of the statement ψ given ϕ can be made as close to 1 as desired, relative to a notion of admissible probability distributions, allowed by prespecified default constraints. On the other hand, a more simple-minded interpretation of the entailment $\phi \mid\!\sim \psi$ – that the conditional probability of ψ given ϕ is greater than or equal to a fixed constant – does not verify them. So at least there is some room for arguing that the laws need not be universal.

We now turn to a somewhat detailed comparison of our information-based semantics with the cumulative and preferential model semantics of Kraus, Lehmann, and Magidor. This is especially interesting since cumulative models give a model-theoretic representation of cumulative consequence relations – those relations satisfying all but the Or law (6).

We recall the basic semantic framework of KLM. Their logical language is that of propositional logic. Formulas are defined as usual with the standard connectives. (For the purely cumulative models, this assumption is not necessary.) Then, one considers a set $\mathcal{U}$ of *worlds*. This may be taken as a subset of the set of all truth assignments to the formulas of propositional logic. Satisfaction of a formula ϕ by a world $u \in \mathcal{U}$ is defined in the standard way. (There is a difference here immediately with our partial model semantics, but perhaps not the crucial distinction.)

Definition 6.1 (cumulative model theory).

- A *cumulative model* is a triple $W = \langle S, l, \prec \rangle$ where S is a set of *states*, $l : S \to 2^{\mathcal{U}}$ is an assignment of worlds to each state, and $\prec$ is a binary relation on S. This relation must satisfy a *smoothness* condition to be given below.
- A state s is said to satisfy a formula ϕ iff for every $u \in l(s)$, $u \models \phi$. We write $s \models\!\!\models \phi$. By $\hat{\phi}$ we mean $\{s \mid s \models\!\!\models \phi\}$.
- The relation $\prec$ is *smooth* iff for every formula ϕ, every element s in $\hat{\phi}$ is either minimal itself in $\hat{\phi}$, or there is a minimal element t of $\hat{\phi}$ with $t \prec s$. (We say that t is minimal if there is no u in $\hat{\phi}$ with $u \prec t$.)

States in S are unspecified, but can be thought of as mental states of an agent. The worlds assigned by l to a state are those worlds which an agent considers possible in the state. The relation $\prec$ between states is the relation which describes preference. If $s \prec t$, then we intend that s is a world which is less exceptional than t. In the standard example, it would be a state in which Tweety is a bird and flies, as opposed to one where Tweety is a bird and does not fly. The smoothness condition looks at first like a technical condition to ensure minimal models of formulas or perhaps theories. However, as KLM point out, it is crucial to verify cautious monotony.

Definition 6.2. Given a cumulative model W, the consequence relation $\mid\!\sim_W$ determined by W is defined by: $\phi \mid\!\sim_W \psi$ iff for any s minimal in $\hat{\phi}$, we have $s \models\!\!\models \psi$.

KLM show that a cumulative model always has a consequence relation satisfying Cut and Cautious Monotony (and conversely).

The definition of cumulative model is quite reminiscent of our default consequence relation, and intuitively is trying to say the same thing. For us, a state is just a situation, and the worlds we consider possible in a situation are the extensions of that situation. Our relation of $\prec$ is just strict subset inclusion. This means that for us, a preferred state for a formula ϕ is one which affirms ϕ, but otherwise has no information about any other properties. Our nonmonotonic consequence relation does not select situations minimal in the set of situations *all of whose extensions* satisfy ϕ, but ones which minimally satisfy ϕ directly. In other words, we could have defined our nonmonotonic consequence relation as follows.

Definition 6.3. Given a default information structure $\underline{A} = A(M, C)$, the "cumulative" consequence relation $\mid\!\sim^c_A$ determined by $\underline{A}$ is defined by the condition $\phi \mid\!\sim^c_A \psi$ iff: for any s minimal in the set

$$\{t \mid (\forall t')(t\, \delta_A\, t' \Rightarrow t' \models \phi)\},$$

we have that all extensions of s satisfy ψ.

It turns out that in the first-order case, the strict subset inclusion relation is not always smooth. This is related to the fact that the set $\hat{\phi}$ is not always Scott open even though ϕ is positive. On the other hand, in finite structures we always get smoothness. So in these cases, we can use default models to get a consequence relation satisfying Cut and Cautious Monotony. Why not then use such a definition? One reason is that we want to use a definition that sanctions reasoning by cases, which the revised definition does not allow. In other words, there is no natural "preferential model", a cumulative model in which the label of a state is a single world.

A more intuitive reason not to use the revised definition is that we think a minimal normal state all of whose extensions satisfy a formula ought itself to satisfy the formula. To see this, recall the counterexample to Cautious Monotony above. The information systems is generated by $\{p\} \vdash \overline{f}$ and $\{f, \overline{f}\} \vdash \emptyset$, and for defaults we take

$$\frac{: f}{f}.$$

As above, we have **true** $\mid\!\sim f$ and **true** $\mid\!\sim f \vee p$, but $f \vee p$ does not $\mid\!\sim f$.

Notice that the *empty* situation is a minimal cumulative "normal world" for $p \vee f$, because all of its extension satisfy f, but of course the empty situation is not a model of $p \vee f$.

Summing up, we might introduce some modal notation to make clear the distinction between our notion and that of KLM. For a formula ϕ, and structure $\underline{A}$, denote by $\mathsf{B}\phi$ the set of situations all of whose extensions positively satisfy ϕ, and by $\mathsf{C}\phi$ the set of situations which directly satisfy ϕ positively. (C stands for "confirm".) Then, our definition of nonmonotonic entailment is $\phi \mid\!\sim_A \psi$ iff

$$\mathrm{Min}(\mathsf{C}\phi) \subseteq \mathsf{B}\psi,$$

while the KLM-style "cumulative" definition would read

$$\mathrm{Min}(\mathsf{B}\phi) \subseteq \mathsf{B}\psi,$$

where Min applied to a collection of situations picks out the collection of subset-minimal elements. So the first definition reads that minimal confirming situations for ϕ support belief in ψ, while the second reads that minimal situations supporting belief in ϕ also support belief in ψ. This shows that our relation is fairly conservative, and it brings out the reliance in our definition on having evidence. We hope that even though we have chosen a strict notion of nonmonotonic entailment, that our positive results gain some added force. That is, we can reason by cases, and we can cut formulas as long as everything in a disjunction is disjoint.

As a final comparison to other work in the literature, our counterexamples to the KLM laws are similar to those of Makinson [20], showing that default logic is not cumulative. There are in fact other options for reclaiming cautious monotony. For example, we could define an order-theoretic version of Brewka's *cumulative default logic* in [6], which was a response to Makinson's examples. Another possibility for the normal default case would be to use a different kind of "extension". In [25], we have introduced such an operator, called a *dilation*. We leave the investigation of the KLM laws under this fixed-point operator to future work.

7. Conclusion

In assessing the novelty of our approach in general, one could claim that generalizing Reiter's ideas to Scott domains is relatively straightforward. We would only argue that it is reasonably difficult to see this given just Reiter's logic as an example. It seems to us that in comparing the major nonmonotonic formalisms, Reiter's notion of default generalizes most easily to an order-theoretic notion. Circumscription is perhaps another one which could be generalized, because looking at minimal models is another order-theoretic idea. It was, however, less clear to us how to do this generalization than it was in the case of defaults.

Our results on open defaults do not completely satisfy us. Kaminski, in [12], has given a full study of the problems with these systems. We need to understand, perhaps, how an open default of the types studied by Kaminski can be interpreted model-theoretically in a default system. For the case of conjunctive defaults we hope to have made some contribution, as an application of our general methodology.

Our definition of nonmonotonic entailment seems to be justified on the grounds of belief structures based on using extensions as accessibility relations. Given a situation, we would be justified in believing a formula if it held in all extensions of that situation. Reiter's notion of "default consequence", for example, is one instance of such a notion, where "situation" means a theory W, and extensions are the usual default extensions in default logic. In this case, notice that "satisfaction" is just given by membership. A formula is a default consequence of W if it is a member of all extensions of W. Notice also that in the first-order case, we get results similar to those provided by circumscription, which by Shoham's results can be seen to be essentially a model-theoretic notion. But once again our models generalize to many more than first-order ones.

The laws we obtain for nonmonotonic entailment are more conservative than the KLM laws. One might see this as evidence that our theory is incorrect. We have indicated that there are other studies (for example, [7]) in which various of the laws are not true. We hope that our counterexamples are realistic enough to at least suggest more careful study of the notion. We see the failure of the laws as pointing out in more detail the differences between defaults as prescriptive devices, as in default logic, and as descriptive ones, as in the probabilistic work. This distinction was very ably pointed out by Pearl in our quotation above.

The potential uses of a domain-theoretic approach to default reasoning still are largely unexplored. Our methodology in this paper led us to ask if there were a domain-theoretic model theory for first-order logic, leading to the information systems described in section 3. This partial model theory seems to be new, independently of nonmonotonic considerations. The generalization of nonmonotonic consequence relations to the Scott topology is also new. All of our positive results on these relations hold in general Scott domains. The first-order systems are just a special case.

Although we make use of one domain-theoretic notion, other methods from domain theory remain to be considered. We would like, for example, to have default versions of function space constructions, and other type-forming operations. Exactly how to go about this is still unclear at present. Other topologies and logics could be considered; for example, the Lawson topology [15], or the well-known intuitionistic semantics using supervaluations over the information inclusion ordering.

We can also consider specializing first-order systems. This could provide a tool for adding defaults to terminological logic systems, another project currently under investigation. Even the special case of feature logic [13] is interesting.

Finally, one of the most exciting prospects for our approach is to combine reasoning about programs with default reasoning, using the common framework of domain theory, which at the very least has shown itself as s superb tool for understanding the former. We would like to provide some evidence that reasoning about actions in AI can be profitably attacked with a common semantic framework.

References

[1] E. Adams, *The Logic of Conditionals* (D. Reidel, Netherlands, 1975).

[2] F. Baader and B. Hollunder, Embedding defaults into terminological knowledge representation formalisms, in: *Proceedings of Third Annual Conference on Knowledge Representation* (Morgan-Kaufmann, 1992).

[3] F. Bacchus, A. Grove, J. Halpern and D. Koller, Statistical foundations for default reasoning, in: *Proceedings of IJCAI* (1993).

[4] J. Barwise, *The Situation in Logic*, 17 (Center for Study of Language and Information, Stanford, California, 1989).

[5] S. Blamey, Partial logic, in: *Handbook of Philosophical Logic*, Vol. III, eds. Gabbay and Guenthner (Reidel, Dordrecht, 1986).

[6] G. Brewka, Cumulative default logic: In defense of nonmonotonic inference rules, *Artificial Intelligence* 50(1) (1991) 183–205.

[7] J. Dix, Classifying semantics of logic programs, in: *Proceedings of First International Workshop on Logic Programming and Non-Monotonic Reasoning*, eds. A. Nerode, W. Marek and V. Subrahmanian (MIT Press, 1991) pp. 166–180.

[8] P. Doherty and W. Lukaszewicz, Distinguishing between facts and default assumptions, in: *Non-Monotonic Reasoning and Partial Semantics*, chapter 3 (Ellis Horwood, 1992).

[9] D.W. Etherington, *Reasoning with Incomplete Information*, Research Notes in Artificial Intelligence (Morgan-Kaufmann, 1988).

[10] R. Guerreiro and M. Casanova, An alternative semantics for default logic (preprint) (1990).

[11] J. Halpern and Y. Moses, A guide to modal logics of knowledge and belief, in: *Proc. IJCAI-85* (1985) pp. 480–490.

[12] M. Kaminski, A comparative study of open defaults. Presentation at 1994 Symposium on Logic and Artificial Intelligence, Fort Lauderdale (January 1994).

[13] R. Kasper and W. Rounds, The logic of unification in grammar, in: *Linguistic and Philosophy* (1991).

[14] S. Kraus, D. Lehmann and M. Magidor, Nonmonotonic reasoning, preferential models, and cumulative logics, *Artificial Intelligence* 44 (1990) 167–207.

[15] J. Lawson, The duality of continuous posets, *Houston Journal of Mathematics* 5 (1979) 357–394.

[16] H.J. Levesque, All I know: A study in autoepistemic logic, *Artificial Intelligence* 42 (1990) 263–309.

[17] V. Lifschitz, On open defaults, in: *Proceedings of the Symposium on Computational Logics* (1990).

[18] V. Lifschitz, Nonmonotonic databases and epistemic queries, in: *Proc. IJCAI-91* (1991) pp. 381–386.
[19] F. Lin and Y. Shoham, Epistemic semantics for fixed-point nonmonotonic logics, in: *Proc. Third Conference on Theoretical Aspects of Reasoning about Knowledge* (1990) pp. 184–198.
[20] D. Makinson, General theory of cumulative inference, in: *Proceedings of Second International Conference on Nonmonotonic Reasoning* (*LNAI 346*) (Springer-Verlag, 1989) pp. 1–18.
[21] J. Pearl, From Adams' conditional to default expressions, causal conditionals, and counterfactuals, in: *Festschrift for Ernest Adams* (Cambridge University Press, 1993). To appear.
[22] R. Reiter, A logic for default reasoning, *Artificial Intelligence* 13 (1980) 81–132.
[23] G.-Q. Zhang, W. Rounds and C. Huang, in: *Proceedings of the 30th Hawaii International Conference on System Sciences*, Maui, Hawaii (1997) Vol. V, pp. 383–391.
[24] W. Rounds and G.-Q. Zhang, Attunement to constraints in nonmonotonic reasoning, in: *Logic Language, and Computation*, Vol. 1, CSLI Lecture Notes, Vol. 58 (CSLI Publication, Stanford, CA, 1996) pp. 479–494.
[25] W. Rounds and G.-Q. Zhang, Domain theory meets default logic, *Logic and Computation* 5 (1995) 1–25.
[26] G. Schwarz and M. Truszczynski, Modal logic s4f and the minimal knowledge paradigm, in: *Proc. Fourth Conference on Theoretical Aspects of Reasoning about Knowledge*, ed. Y. Moses (1992) pp. 184–198.
[27] D.S. Scott, Domains for denotational semantics, in: *Lecture Notes in Computer Science 140* (1982).
[28] Y. Shoham, A semantical approach to nonmonotonic logics, in: *Readings in Nonmonotonic Reasoning*, ed. M.L. Ginsberg (Morgan-Kaufmann, 1987).
[29] S. Vickers, *Topology via Logic* (Cambridge University Press, 1988).
[30] M. Young, Nonmonotonic sorts for feature structures, in: *Proceedings of AAAI-92* (1992) pp. 596–601.
[31] G.-Q. Zhang, *Logic of Domains* (Birkhauser, Boston, 1991).