Research Report AI–1994–10

Nonmonotonic Consequence in Default Model Theory

Guo-Qiang Zhang
Department of Computer Science
University of Georgia

William C. Rounds
Artificial Intelligence Laboratory
University of Michigan

NONMONOTONIC CONSEQUENCES IN DEFAULT MODEL THEORY [1]

Guo-Qiang Zhang
Department of Computer Science
University of Georgia
Athens, Georgia 30602
`gqz@cs.uga.edu`
and
William C. Rounds
Artificial Intelligence Laboratory
University of Michigan
Ann Arbor, Michigan 48109
`rounds@engin.umich.edu`
December 7, 1994

Abstract. Default model theory is a nonmonotonic formalism for representing and reasoning about commonsense knowledge. Although this theory is motivated by ideas in Reiter's work on default logic, it is a very different, in some sense dual framework. We make Reiter's default extension operator into a constructive method of building *models*, not theories. Domain theory, which is a well established tool for partial information in the semantics of programming languages, is adopted as the basis for constructing *partial* models. One of the direct advantages of default model theory is that nonmonotonic reasoning can be conducted with *monotonic logics*, by using the method of *model checking*, instead of *theorem proving*.

This paper reconsiders some of the laws of nonmonotonic consequence, due to Gabbay and to Kraus, Lehmann, and Magidor, in the light of default model theory. We remark that in general, Gabbay's law of cautious monotony is open to question. We consider an axiomatization of the nonmonotonic consequence relation omitting this law. We prove a representation theorem showing that such relations are in one to one correspondence with the consequence relations determined by extensions in Scott domains augmented with default sets. This means that defaults are very expressive: they can, in a sense, represent any reasonable nonmonotonic entailment. Results about what kind of defaults determine cautious monotony are also discussed. In particular, we show that the property of unique extension guarantees cautious monotony, and we characterize default structures which determine unique extensions.

[1] An earlier version was presented at the Third Bar-Ilan Symposium on the Foundations of Artificial Intelligence, Bar-Ilan University, Israel, July 1993.

1 Introduction: Default models versus theories

The purpose of this paper is to explore just one of the directions suggested by making default systems, as in Reiter's work [19], into constructions for building partial models for various kinds of logic. Our reasons for undertaking this task are not just to propose another nonmonotonic formalism, but to try to correct some of the well-known problems with default logic, while at the same time generalizing default systems to a far wider setting than just first-order logic. This paper does not address the former issue in any detail, but it provides some evidence that the latter project is a good one. For the sake of completeness, though, we recall some of the difficulties with default logic, before giving a general introduction to our methodology, which we call *default model theory*. We presume a basic familiarity with default logic on the reader's part; for a general overview consult [4].

1.1 Why not default logic?

Among the recognized problems with default logic are the following.

- There can be multiple extensions, or competing theories.
- Extensions may not exist.
- Even when extensions do exist, it can be too costly or impossible (in terms of an effective procedure) to find them, because consistency for first order logic is undecidable.
- Reasoning by cases fails.
- Default logic is not cumulative [16].

Proposals for amending these faults are numerous, such as disjunctive default systems [18], Brewka's cumulative default logic [5], and so forth. However, there is no known remedy so far that fixes these problems all at once. On the contrary, it is often the case that further problems are introduced with the amendments.

We do not propose here to solve all of these problems with one magic system, but we do want to suggest one reason why so many approaches have not worked well. This is because nonmonotonic reasoning is essentially about incomplete information (or partial information), *in a model-theoretic sense*, but the model theory of the logics used is not. Incomplete information is a key aspect of nonmonotonic reasoning because you conclude, for example, that a specific bird flies *in absence of the information* that it is a penguin. You expect to meet a flight at the airport at its scheduled arrival time *in absence of the information* that the flight is delayed. You base your decisions (in the second case especially) not only on properties assumed about airports, but on the partiality inherent in a real-world situation.

Many current approaches to default reasoning are based on the total models used in classical first order logic [2]. This presents a fundamental mismatch between the theoretical

[2]Of course, probabilistic tools can and have been used for this purpose. But in many cases, probability distributions and/or statistical information is unavailable.

tool on the one hand, and the phenomena we want to describe on the other. The basic view of classical logic is total: every issue is settled as either *true* or *false*, and truth values never change. In commonsense reasoning, our conclusions cannot be supported by proofs in the mathematical sense, and we need to take action in situations of partial information. The only place such information resides, in traditional default logic, is in the incompleteness of default theories. But it seems that the lack of information about an airplane's arrival time is not well captured by an incomplete scientific theory of its particular flight.

One can argue that the reason for using first order logic is that it is the most well established formalism that is familiar to many. But familiarity with a tool should not be enough reason for applying it in every knowledge representation task. To be sure, a case can be made in favor of using classical logic: Any meta-formalism withstanding the scrutiny of mathematical rigor may eventually be translated into certain forms of mathematical logic, first or higher order (such as ZF set theory). Therefore, the use of classical logic is in some sense inevitable. So, why not start there right from the beginning? The answer is clear: one should use mathematical tools appropriate for a given modeling task. Differential equations, for example, are entirely appropriate for predicting trajectories. First-order logic, including the theory of real numbers implicit in *Principia Mathematica*, is an inappropriate tool for that purpose.

The first systematic and extensive research project questioning the ubiquity of first-order logic, and in fact formal logic in general for the semantics of natural language, is perhaps Situation Theory, pioneered by Barwise and Perry [3] in the early 80's. Situation Theory takes the stand that the key issue for (any) logic is first and foremost its informational character, not languages or proof systems.

We recognize that stressing the need for an independent basis for nonmonotonic reasoning, which treats partial information as a first class citizen, should not be taken as a denial of the value of more than a decade's research in this area. We believe that many of the intuitions are sound and productive, and we want to take advantage of them. The idea of an extension in default logic, for example, is one of the key methods for building partial models in default model theory.

1.2 Why default model theory?

Is there a way to get rid of the problems with default logic, *once and for all*? Although more work still needs to be done, we believe that default model theory is on the right track towards this goal.

The basic observation of default model theory is that nonmonotonic reasoning is fundamentally about partial information, and, therefore, needs to be explained using a full-fledged theory of partial information. If we should not automatically build everything by extending first order logic, then what else can we use? Fortunately, results in the area of semantics of programming languages provide a rich class of tools for dealing with partial information. Domain theory, as developed by Scott and many others since the early 70's, is concerned with computable or constructive methods for building partial models of programs. A number of tools have been developed in domain theory for dealing with a variety of partial models, nondeterministic programs, Petri nets, event structures, and the like. Partial information is a core concept in this theory.

For this reason, our basic approach to default reasoning uses the theory of domains. We view Reiter's default rule systems as nondeterministic 'algorithms' for populating a (partial) model with elementary facts about particular situations. This is in exact analogy with (in fact a generalization of) Scott's concrete characterization [23] of consistently complete ω-algebraic partial orders (Scott domains) using the technique of information systems. In that construction, elementary 'tokens' are added to a partial model according to a monotonic rule system. We augment Scott's monotonic rules with default rules to obtain a class of default information systems. For these systems, we construct extensions of a given world, as in Reiter's work [19], but these extensions are models, not theories.

We suggest that Reiter's default systems should be regarded as *semantic*, not syntactic notions. This is a radical reconstruction of default reasoning. We call our thesis, together with its supporting evidence, *default model theory.* As a result, many of the current applications of nonmonotonic logic can be seen, using a unified framework, in an entirely different light. We provide an introduction to the basic properties of default structures in [21]. We have proposed a default partial model semantics for first order logic in [22]. Soundness and completeness results for a propositional modal logic of belief have been obtained, again using default structures as models [20].

We emphasize that the purpose of default model theory is *not to amend or extend default logic, but to parallel it.* Therefore, we suggest the reader not to try to fit this paper into the realm of default logic.

At a first glance, the shift from the proof theory of default logic to a semantic theory about partial information may not seem too drastic. However, the philosophy behind default model theory is fundamentally different, and the consequences are significant. We now explain why, in default model theory, at least some of the problems with default logic can be ameliorated.

Multiple extensions. When extensions are regarded as partial (possible) worlds, the extension relation is similar to the accessibility relation in Kripke structures. A default rule functions here not as an extended proof rule, but as part of a constructive procedure for building an agent's most plausible worlds extending the current one. There can be many different worlds considered most plausible from the current world. So the possibility of multiple extensions becomes a feature, not a bug – Kripke structures would be rather uninteresting if there were only one world accessible from the current one.

Existence of extensions. When defaults are regarded as a constructive method for building worlds, we make some room for accommodating different model building procedures. Reiter's extension operator, when phrased model-theoretically, remains one of the key 'algorithms' for building preferred worlds. However, extensions may not exist in some reasonable cases. To cope with this, we have introduced the notion of a 'dilation' [21], which is a robust generalization of the notion of an extension. Dilations exist in all reasonable cases, including semi-normal defaults. Extensions, when exist, are minimal dilations. Even dilations may not exist for some defaults, however. But examples [21] show that when that happens, the blame goes to the inherent incompatibility of defaults, not to the method of dilation itself. One can imagine having two defaults

$$\frac{X : \emptyset}{a}, \frac{X : \emptyset}{\neg a},$$

which are fundamentally incoherent, with the 'justification' parts unrelated to the applicability of the defaults. One should not expect any notion of an 'extension' to exist for such a default set. (This paper does not treat dilations *per se*. We are concerned with the standard case of normal defaults.)

Monotonic logics. One direct benefit of making default models and not default theories is as follows. Our modal logics for belief, derived from default semantics, are *monotonic* and they have all the properties one expects of a logic in the standard sense. For example, we have a sound and complete proof system for a propositional modal logic, including the belief modality B, where $\mathsf{B}\varphi$ means that in our current world, φ holds in all default extensions. Our modal logic is decidable, with validity being co-NP-complete.

What is it that makes it possible to do *nonmonotonic* reasoning using a *monotonic* logic? The fact that extensions in default model theory are partial models makes it possible to reason by *model checking*, as advocated by Halpern and Vardi [10]. Nonmonotonic reasoning can be reduced to, for example, checking if $x \models \mathsf{B}\varphi$ holds, instead of whether or not φ follows from x by applying certain extended deduction rules for first order logic. A simple example illustrates the difference here: in partial models, it may be the case that Fred's deadness is not supported by a model which only has the information that Tweety is a bird. Using classical negation for the sake of argument, this would be expressed by

$$\{\texttt{Bird}(\texttt{Tweety})\} \models \neg\texttt{Dead}(\texttt{Fred}).$$

But it is absurd to assert that $\texttt{Bird}(\texttt{Tweety})$ should have $\neg\texttt{Dead}(\texttt{Fred})$ even as a default consequence. In default model theory, nonmonotonicity appears in a different form: a belief $\mathsf{B}\varphi$ supported in a situation x (written $x \models \mathsf{B}\varphi$) may become unsupported in a situation y with more information ($x \subseteq y$, but $y \not\models \mathsf{B}\varphi$).

Computational cost. Default model theory is intuitively computationally less complex than the highly undecidable default logic. This is partly because model checking is generally easier than theorem proving [10], and partly because situations are typically made of only atomic, propositional statements (considered as databases), not arbitrary first order formulas.

Here we share the view of Levesque [13] about vivid knowledge bases:

> In the early days of AI, we had a model of intelligence based on powerful general methods such as resolution, heuristic search, means-end analysis, and so on. The end result of this view of AI was inevitably combinatorial explosions and disaster. Now, in the enlightened era of knowledge-based AI, instead of general methods, we use large amounts of domain-dependent knowledge, and instead of combinatorial explosions, we end up with expert systems that work.

But

> Why should it be computationally easy to apply large amounts of domain-dependent knowledge?

The reason is, as Levesque points out, that our knowledge bases are vivid, in the sense that it is conjunctive and propositional in nature, and it can be regarded as a model of itself. Our situations x are just like vivid knowledge bases. These arguments suggest that it will be computationally feasible to do nonmonotonic reasoning in default model theory.

Reasoning by case analysis. Suppose that we are given the defaults

$$\frac{q : p}{p} \text{ and } \frac{v : p}{p}.$$

Now we would like, given $q \vee v$, to be able to conclude 'by default' that p. Using default logic we cannot even apply the defaults, because neither q nor v is provable from $q \vee v$. Under default model theory, because of the model checking method used, any state (partial model) x which supports $q \vee v$ does have p in all its extensions. This is because whenever $x \models q \vee v$, it must be the case that either $x \models q$, or $x \models v$. It then turns out that in either case p is in the relevant extension.

The use of default models is in some sense anticipated by the work of Guerreiro and Casanova [7] and Lifschitz [15]. Our definition is much more radical than theirs. Although their semantics involves a fixed-point construction in model-theoretic space, they still regard extensions as theories. 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.

1.3 Nonmonotonic entailment

We now focus on the main topic of the paper: the laws that can govern nonmonotonic reasoning. Since Gabbay [6], a considerable amount of work has been done on axiomatizations of nonmonotonic consequence relations, here denoted by $\mid\sim$. Before we get into the more technical details of our framework, we give an intuitive description of our interpretation of $\mid\sim$. Let φ and ψ be certain formulas, propositional or otherwise. With respect to a given default structure, a formula φ nonmonotonically entails a formula ψ, written as

$$\varphi \mid\sim \psi,$$

iff ψ is satisfied in every extension of any informationally minimal partial model of φ.

Several points should be made about our definition of nonmonotonic consequence. First, it uses extensions as models, not theories. Thus it makes sense in default model theory to say that a formula is satisfied in an extension. Second, the intuitive interpretation of our entailment is "if I can confirm only the information given by φ, then I can believe (skeptically) ψ." This idea is related to, but is not the same as, the notion of "minimal knowledge", or "only knowing". [9, 14]. These authors initiated the use of *maximal models* to capture "minimal knowledge", because for them partiality of information is captured by the partiality of a theory. We capture the notion in a dual sense: by the use of minimal partial models. We regard partial models as sets of tokens, each of which carries a "unit" of information. Minimality of a partial model is measured by set inclusion. If, for example, we are given the formula $Bird(tweety)$, then the minimal information conveyed by the formula is a partial model of the formula simply consisting of the one 'first-order tuple' $\langle\langle bird, tw; 1\rangle\rangle$. Given a theory I, we look at all the minimal partial models of the formulas in I. We call this the 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.

What properties should be required of $\mid\!\sim$? Many authors, in particular Kraus, Lehmann, and Magidor [11], have considered Gabbay's axiom of *cautious monotony:*

$$X \mid\!\sim a \;\&\; X \mid\!\sim b \Rightarrow X, a \mid\!\sim b.$$

This axiom is often interpreted as (for example, in [12]): the learning of a new fact, the truth of which was expected, cannot invalidate a previous conclusion.

Following our intuitive interpretation of nonmonotonic entailment, we soon realize that cautious monotony need not be universally valid. Here is an example which seem to invalidate the axiom. It relies on the canonical example of birds, penguins and flying. Any notion of nonmonotonic consequence will have to deal with this example, in such a way that

$$\texttt{bird} \;\mid\!\sim\; \texttt{fly} \tag{1}$$

holds in the system. By a standard logical weakening, we should also have

$$\texttt{bird} \;\mid\!\sim\; \texttt{fly} \;\vee\; \texttt{penguin}. \tag{2}$$

Applying cautious monotony, we get

$$\texttt{bird},\;\; (\texttt{fly} \;\vee\; \texttt{penguin}) \;\mid\!\sim\; \texttt{fly}. \tag{3}$$

But it is questionable to accept (3) as a reasonable instance of nonmonotonic entailment. The natural language reading of (3) is something like this: "a bird flies, even if it is a penguin", which is counter-intuitive.

This of course does not mean that there is no other interpretation which could justify the cautious monotony law. In fact, probabilistic analyses of defaults, such as those in Adams [2], and subsequently Pearl [17], among many others, have included cautious monotony as one of the 'core laws' which should hold in any reasonable calculus of approximate reasoning.

Statement (3), for example, can be justified on a probabilistic basis, since the non-flying birds are very rare compared to the bird population. 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 [17, Section 1].

We agree in part with this analysis, but we also note that generics have been analyzed in the tradition of philosophical logic as if they were intended to convey at least approximate information about the world.

Adams' ϵ-semantics verifies the cautious monotony law, but involves interpreting defaults in a limit sense: $\varphi \mid\!\sim \psi$ 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 $\varphi \mid\!\sim \psi$ – 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 cautious monotony needs not be universal.

We now turn to a brief comparison of our information-based semantics with the cumulative model semantics of Kraus, Lehmann, and Magidor. This would explain, in particular, why cautious monotony holds in their framework, but not ours. Intuitively, a cumulative model is made of a set of states, a possible worlds function, which assigns to each state a set of plausible worlds, and a well-founded preference relation between states. With respect to a cumulative model, $\varphi \mid\!\sim \psi$ iff for any most (maximally) preferred state all of whose possible worlds satisfy φ, the plausible worlds in that state also satisfy ψ. Moreover, the *preferential* models are those where there is exactly one world plausible in each state. Cautious monotony now follows fairly easily in both these cases.

It seems natural in our case to take the inclusion relation on partial models as the preference relation, with minimality of information corresponding to maximally preferred states. (This gets at the idea of "all else being equal.") Then the worlds plausible in a situation correspond to the extensions of that situation; this is exactly the tradition of Kripke semantics, except that the extensions are explicitly constructed using defaults. With respect to a default model, $\varphi \mid\!\sim \psi$ iff given any state x with only the minimal information to support φ, all the extensions for x satisfy ψ. We have, for example, $\texttt{bird} \mid\!\sim \texttt{fly}$, because given a state where you only have the information that something is a bird (and without the extra information that it is a penguin), the most plausible states support that it flies. This matches our intuition: to conclude a bird flies in absence of the information that it doesn't.

It is now clear that we cannot accept

$$\texttt{bird},\ \ (\texttt{fly}\ \vee\ \texttt{penguin})\ \mid\!\sim\ \texttt{fly}$$

as a valid instance of nonmonotonic entailment under our interpretation. There are two minimal information states which support `bird` $\wedge$(`fly` $\vee$ `penguin`), one of which is the state with only the information that the bird is a penguin. Of course its extended states do not support that the penguin flies.

1.4 Contributions of this paper

One cause for the seeming failure of cautious monotony is the disjunction used in the setting where pieces of information have propositional structure (call this the logic setting). This may lead us to believe that a similar failure would not occur in a setting where nothing is assumed about internal structure of pieces of information (call this the abstract setting.) However, things are not that simple. As noted earlier, the set of axioms satisfied by the induced nonmonotonic consequence relation crucially depends on the method used for building the partial worlds. Although Reiter's notion of an extension is an important one, it violates cautious monotony even in the abstract setting. (However, we do not have a strong intuition why cautious monotony should fail in this case. In fact, it need not fail in the abstract setting – see discussion in the conclusion section.)

The rest of this paper will focus on the abstract setting of nonmonotonic entailment. We summarize the main results of this paper, in order of perceived importance:

- We recast the issues involved in the study of the abstract nonmonotonic entailment relation into the bigger picture of default model theory. This paper is not merely about "yet another theory for nonmonotonic entailment". Rather, it is an integral part of a more ambitious effort to reformulate nonmonotonic reasoning, as explained earlier. The main tool used to do this is the notion of a "nonmonotonic information system". This is a generalization of Scott's information systems, so we are already including all Scott domains in our theory.

- We prove a representation theorem which characterizes the properties of a nonmonotonic entailment relation defined via the axioms of a nonmonotonic information system, in terms of default rules and extensions. This establishes a close connection between defaults and the abstract nonmonotonic entailment relation, which has not been established in such a general setting (most representation theorems assume a framework of a classical propositional language).

- The abstract nonmonotonic entailment relation derived by using the construction of extensions does not in general satisfy the cautious monotony law. Examples and discussions are provided to explain this.

- Subclasses of defaults are identified for which cautious monotony does hold. In particular, unique extension guarantees cautious monotony, and we provide a sufficient condition for defaults to determine unique extension.

- The abstract nonmonotonic entailment relation is represented in the setting of complete partial orders and Scott open sets, showing the generality of the issues involved. It also establishes a framework where the rich set of results in domain theory can be directly applied.

- Finally, a distinguished aspect of our general methodology is that it is *semantics based*. This means that we do not start from a pre-selected set of axioms for nonmonotonic entailment, and then think about what kind of models are appropriate. Instead, we start from the default models right at the beginning, and then derive a set of valid axioms for the models.

In the sequel, we give a brief introduction to Scott domains in Section 2, mostly via the notion of information systems. The basics of default model theory are presented in Section 3, and our representation result is given in Section 4, for non-cumulative systems. Section 5 studies subclasses of defaults for which cautious monotony holds. Section 6 discusses nonmonotonic entailment in a general topological setting, a natural step from the domain theory point of view. The last section discusses further work.

2 Domain Theory, Information Systems, and Cpos

This section gives a brief introduction to domain theory for readers who may not be familiar with this area.

Domain theory is a branch of theoretical computer science developed by Scott and others for the semantics of programming languages. The development of domain theory, in the late 1960's, started with the observation that there were no mathematical models for the lambda calculus, although the calculus was taken as a formalism in which to interpret programming languages such as Algol 60. To have a sound foundation for the semantics of programming languages, a mathematical model for the lambda calculus seemed to be desirable. However, it was not easy to come up with such a model, because of the high-order functional characteristic of the lambda calculus: a term can take another term as its argument. There had to be a way to solve equations like $(D \to D) = D$, where $(D \to D)$ is a set of certain functions from D to D. A naive set theoretical construction, taking $(D \to D)$ to be that set of *all* functions from D to D, does not work, because $(D \to D)$ always has a larger cardinality than D for non-empty D. Scott's idea was to work with partial objects, instead of total objects. Total objects, however, can be approximated by increasingly better approximations. This leads us to the theory of domains.

The basic structures of domain theory are complete partial orders (cpos). A complete partial order is a partial order $(D, \sqsubseteq)$ with a least element $\perp$, and least upper bounds of increasing chains

$$x_0 \sqsubseteq x_1 \sqsubseteq x_2 \sqsubseteq \cdots.$$

When we write $x \sqsubseteq y$, we mean that x is an approximation to y, or y contains more information than x. A subset $X \subseteq D$ is bounded (or compatible, consistent, $X \uparrow$) if it has an upper bound in D. Thus for a bounded set X all the elements in X can be thought of as approximations to a single element. A compact (or finite) element a of D is one such that whenever $a \sqsubseteq \bigsqcup_i x_i$ with

$$x_0 \sqsubseteq x_1 \sqsubseteq x_2 \sqsubseteq \cdots,$$

we also have $a \sqsubseteq x_k$. 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 which is the least upper bound of a increasing chain 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.

There is much more to domain theory than can be summarized here. Readers who are interested in finding more about it should consult [23, 8, 25, 1, 26]. In the remainder of this section we review a concrete representation of Scott domains, called information systems, which will be used in this paper.

An information system consists of a countable set A of tokens, a subset *Con* of the set of finite subsets of A, denoted as $\mathrm{Fin}(A)$, and a relation $\vdash$ between *Con* and A. The subset *Con* on A is often called the consistency predicate, or the coherent (compatible) sets. The relation $\vdash$ is called the entailment relation, or the background constraint. The intended usage of the consistency predicate and the entailment relation suggests that we put some reasonable conditions on them. This results in the following definition.

Definition 2.1 *An information system is a structure* $\underline{A} = (\,A,\ \mathit{Con},\ \vdash\,)$ *where*

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

which satisfy

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

Let us explain these conditions by taking the token set to be $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. The first condition says that any subset of a coherent set must be coherent. So, if $\{1, 2, 3\}$ is a coherent set for an information system, then $\{1, 2\}$ must also be a coherent set, and similarly for $\{2, 3\}$, $\{1, 3\}$, etc. The second condition requires that each single token must be coherent – otherwise we can do without it. The third condition says any token entailed by a finite set must be coherent with the set. For example, if we have $\{1, 2, 3\} \vdash 4$, then $\{1, 2, 3, 4\} \in \mathit{Con}$. The fourth condition is a form of reflexive property for $\vdash$. The last condition is a kind of transitive property for $\vdash$. It makes most sense if we consider a finite coherent set X as a conjunction of tokens in X. That is, when we write $\{1, 2, 3\} \vdash 4$, we can imagine $1 \wedge 2 \wedge 3 \vdash 4$. It is then reasonable to insist that if $1 \wedge 2 \wedge 3 \vdash 4$, $1 \wedge 2 \wedge 3 \vdash 5$, and $4 \wedge 5 \vdash 6$, then $1 \wedge 2 \wedge 3 \vdash 6$.

Although the consistency predicate *Con* and the entailment $\vdash$ can be regarded as a generalization of those derived from the first order logic, it is important to note that they are *primitive* for an information system. Therefore, the decidability of *Con* and $\vdash$ are often *built in* to typical information systems.

It should also be pointed out that the user has full control of what *Con* and $\vdash$ are going to be, as long as the properties required for an information system are not violated. For the same token set A, a change of relations *Con* and $\vdash$ merely results in a different information system. Whether one information system is better than another depends on the application at hand.

Example. This information system is for approximating real numbers. For tokens, take the set A to be pairs of rationals $\langle q, r \rangle$, with $q \leq r$. The idea is that a pair of rationals stands for the information 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 tokens. We do not negate them or disjoin them, *by our choice*. It is straightforward to verify that the five properties for information systems hold for this case.

□

The notion of consistency can be easily extended 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 2.2 *The collection* $|\underline{A}|$ *of* **ideal elements** *of an information system consists of subsets* x *of propositions which are*

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

Information systems provide a way to express knowledge about the world in terms of coherence *Con* and constraints $\vdash$. Elements of an information system are made of tokens which are coherent, and which respect the constraint. They describe situations, or states of knowledge, which are 'partial models' in the sense that situations usually do not settle the truth of every issue. One can also regard an information system as a set of rules from which to construct a Kripke structure. The ideal elements are the possible worlds. The accessibility relation is set inclusion, corresponding to the direction of information increase. Moreover, an atomic token is supported in a world just in case that the world contains the token. In this setting, genuine negative information has to be encoded in tokens as well – we regard falsity as positive information. So in the case where tokens do encode falsity – which they need not always do – truth assignment is a partial function, and we should probably call these structures partial Kripke structures.

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 the empty set, which can be regarded as a special interval $(-\infty, \infty)$.

□

Theorem 2.1 *(Scott) 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.*

This result is basic in domain theory. We mention this result here to convince the reader that information systems are thought of as (generators of) semantic structures, not syntactic entities. (More precisely, our purpose is not to study the syntax of information systems, which seems straightforward.) Of course, one cannot make the distinction between syntax and semantics absolute. Any "constructive" mathematical model must by nature be built from primitive pieces of "syntax".

3 Normal Default Structures

Now we come to the main definitions of default model theory. Normal default structures are information systems extended with a set of 'defaults' expressing an agent's belief. Each default takes the form

$$\frac{X : a}{a},$$

with X a finite consistent set, and a a single token of the underlying information system. The idea of the defaults is that one can generate models (states) by finding X as a subset of tokens in a state under construction, checking that a is consistent with the current state, and then adding the token a.

Definition 3.1 *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 defaults, each element of which is written as $\frac{X : a}{a}$, with $X \in \mathit{Con}$, $a \in A$. If each default is of the form $\frac{: a}{a}$, we call the default structure precondition free.

The notion of deductive closure associated with standard information systems will be often used. Let $(A, \mathit{Con}, \vdash)$ be an information system, and G a coherent subset of A. The *deductive closure* of G is the set

$$\overline{G} := \{a \mid \exists X \subseteq^{fin} G.\ X \vdash a\},$$

where $\subseteq^{fin}$ stands for "finite subset of".

Extensions are a key notion related to a default structure. An extension of a situation x is a partial world y extending x, constructed in such a way that everything in y faithfully reflects an agent's belief expressed by defaults. If the current situation is x, then because it is a partial model, it may not contain enough information to settle an issue (either positively or negatively). Extensions of x are partial models containing at least as much information as x, but the extra information in an extension is only plausible, not factual.

The following definition is just a reformulation, in information-theoretic terms, of Reiter's own notion of extension in default logic.

Definition 3.2 *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*

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

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

Example. Using the 'approximating real' system described earlier, 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 $\frac{: a}{a}$, where a ranges over rational pairs $\langle p, q \rangle$ such that $p \leq 0$ and $q \geq 1$, together with those pairs $\langle r, s \rangle$ such that $r < \pi$ and $s > \pi$.

In the ideal domain, we refer to elements by the real intervals to which they are isomorphic. 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 $[\pi, \pi]$.

□

More example: 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

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

We may take $\vdash$ to be trivial: $X \vdash \langle i, j \rangle$ iff $\langle i, j \rangle \in X$. Extensions are those placements of queens which do not violate any constraints of the rules of chess, and which cannot be augmented without causing a violation. The eight queens problem can be rephrased as finding all the extensions of size 8 for $\emptyset$.

□

We now look at ways extensions can be constructed for normal defaults. Again, our construction is a generalization of Reiter's construction of an extension for a normal default theory.

For a set T of tokens, let $\mathsf{M}(T)$ be the set of maximal consistent subsets of T. For example, if T is itself consistent, then $\mathsf{M}(T) = \{T\}$. Given a normal default information structure $(A, Con, \vdash, \Delta)$, extensions can be constructed in the following way for a given x.

Let $x_0 = x$. For each $i > 0$, let

$$x_i \in \mathsf{M}(\overline{x_{i-1}} \cup \{a \mid \frac{X : a}{a} \in \Delta \ \& \ X \subseteq x_{i-1}\})$$

with $x_i \supseteq \overline{x_{i_1}}$.

It is easy to see that x_i is an increasing chain. Let $M := \bigcup_{i \in \omega} x_i$. We show that

$$M = \bigcup_{i \in \omega} \phi(x, M, i).$$

We prove by mathematical induction that

$$x_i = \phi(x, M, i)$$

for each $i \geq 0$.

The base case is clear. Assume $x_i = \phi(x, M, i)$ for some i. Since

$$\begin{aligned}\overline{x_i} \subseteq x_{i+1} &\in \mathsf{M}(\overline{x_i} \cup \{a \mid \frac{X : a}{a} \in \Delta \ \& \ X \subseteq x_i\}) \\ &= \mathsf{M}(\overline{\phi(x, M, i)} \cup \{a \mid \frac{X : a}{a} \in \Delta \ \& \ X \subseteq \phi(x, M, i)\}),\end{aligned}$$

we have

$$\overline{\phi(x, M, i)} \cup \{a \mid \frac{X : a}{a} \in \Delta \ \& \ X \subseteq \phi(x, M, i) \ \& \ \{a\} \cup M \in Con\} \subseteq x_{i+1}.$$

On the other hand, if $a \in x_{i+1}$, then either $a \in \overline{x_i}$ or $a \in x_{i+1} - \overline{x_i}$. For the latter, we have $\frac{X : a}{a} \in \Delta$, $X \subseteq \phi(x, M, i)$, and $\{a\} \cup M \in Con$. Therefore $a \in \phi(x, M, i+1)$.

Thus $x_{i+1} = \phi(x, M, i+1)$. Note, however, it is not necessary to take a maximal consistent set at each step. As long as maximal consistent sets are taken infinitely often, we will get an extension.

We have, in effect, proved the following generalized Reiter theorem.

Theorem 3.1 *Extensions always exist for normal default structures.*

The question arises as to whether every extension for x can be constructed in this way. Although this is true for precondition free defaults, it is not true in general.

Example. Not all extensions can be constructed using the method described above. Consider the following default structure with defaults

$$\frac{: a}{a}, \ \frac{: b}{b}, \ \frac{a : c}{c},$$

but $\{a, b, c\}$ inconsistent. $\{a, b\}$ is an extension for $\emptyset$ constructed by this method. However, $\{a, c\}$, which is also an extension for $\emptyset$, can not be constructed in this way.

$\square$

Example. This example shows that there are extensions that cannot be constructed by taking any maximal consistent sets. This infinite default structure has defaults

$$\frac{: a_1}{a_1}, \ \frac{a_1 : a_2}{a_2}, \ \frac{a_2 : a_3}{a_3}, \cdots$$

and

$$\frac{: b_1}{b_1}, \ \frac{: b_2}{b_2}, \ \frac{: b_3}{b_3}, \cdots,$$

such that $\{b_i, a_{i+1}\}$ is inconsistent for each $i > 0$. One can check that $y = \{a_i \mid i > 0\}$ is an extension of $\emptyset$, and $\phi(x, y, j) = \{a_i \mid j \geq i > 0\}$. None of the $\phi(x, y, j)$'s is maximal consistent in

$$\overline{\phi(x, y, j-1)} \cup \{a \mid \frac{X : a}{a} \in \Delta \ \& \ X \subseteq \phi(x, y, j-1)\},$$

because $\{b_j\} \cup \{a_i \mid j \geq i > 0\}$ is such a maximal consistent set.

$\square$

We now summarize properties of extensions for normal default structures. (Recall that $x \epsilon y$ means y is an extension of x.)

- Extensions always exist.
- If $x \epsilon y$ then $y \supseteq x$.
- $x \epsilon y$ and $y \epsilon z$ if and only if $y = z$.
- If $x \epsilon y$ and $x \epsilon y'$, then either $y = y'$ or $y \cup y' \notin \mathit{Con}$.
- If $x \epsilon z$ and $y \subseteq z$, then $\overline{x \cup y} \epsilon z$.

The proofs of these facts again follow those of Reiter, so we omit them.

4 Defaults and Nonmonotonic Consequences

We are interested in the relation between default structures and nonmonotonic entailment relations.

In our opinion, axiomatizing nonmonotonic entailment $\mid\!\sim$ without referring to a consistency predicate *Con* does not seem to fully capture our intuition. One thing we would like to rule out, in particular, is an example like

$$a \mid\!\sim \overline{a}$$

where $\overline{a}$ is in some sense the negation of the token a. It does not seem reasonable to entail a piece of information freely from an inconsistent piece. At first sight, this assertion seems to be directly at odds with the intuition expressed by Kraus, Lehmann, and Magidor, in which the authors argue for a statement such as

if I'm the Queen of England then anything goes.

Here, "anything goes" may include "I am not the Queen of England".

Here it is important to remember that our notion of "entailment" is going to be that the consequent is a piece of information supported by all plausible worlds extending a current situation itself supporting the antecedent. It is not reasonable to have such worlds contain pieces of information which contradict facts already known in the current world. Perhaps better said: we will never entertain worlds which contain contradictory facts like a person who is not the Queen of England being the Queen of England.

In the study of nonmonotonic consequences, the following Gentzen-style axioms are often considered. (Here, $X \mid\!\sim Y$ is an abbreviation for $\forall b \in Y\ X \mid\!\sim b$.)

Identity: $a \in X \Rightarrow X \mid\!\sim a$.
Cautious monotony: $X \mid\!\sim a \;\&\; X \mid\!\sim b \Rightarrow X, a \mid\!\sim b$.
Cut: $X \mid\!\sim T \;\&\; T, Y \mid\!\sim b \Rightarrow X, Y \mid\!\sim b$.
Cautious cut: $X \mid\!\sim T \;\&\; T, X \mid\!\sim b \Rightarrow X \mid\!\sim b$.

For these axioms, X, Y, and T range over finite sets of formulas, and a and b are single formulas, and a set X is thought of conjunctively. (The notation X, T stands for $X \cup T$.) We use the form of these axioms as a guide for our abstract formulation.

We next introduce the notion of a nonmonotonic system. Now, in contrast to the above axioms, we let X, Y, and T range over finite subsets of an *arbitrary countable or finite set* A.

Definition 4.1 *A nonmonotonic system is a triple*

$$(A, Con, \mid\!\sim)$$

where Con is a collection of finite subsets X of A, called the consistent or coherent sets; and $\mid\!\sim$ is a subset of $Con \times A$, called the relation of nonmonotonic entailment, which satisfies the following axioms:

1. $X \subseteq Y \in Con \Rightarrow X \in Con$,
2. $a \in A \Rightarrow \{a\} \in Con$,
3. $X \mid\!\sim T \Rightarrow X \cup T \in Con$,
4. $a \in X \;\&\; X \in Con \Rightarrow X \mid\!\sim a$,
5. $X \mid\!\sim T \;\&\; T \cup X \mid\!\sim b \Rightarrow X \mid\!\sim b$.

Note that $X \mid\!\sim T$ is an abbreviation for $(\forall a \in T)(X \mid\!\sim a)$, where T is a finite subset of A.

It is clear that the concept of a a nonmonotonic system is a generalization of the concept of information system mentioned in Section 2. Axioms 1, 2, and 4 are adopted here without change.

Axiom 3 is derivable for information systems, but bears some consideration for nonmonotonic entailment. It can be regarded as a very weak form of the cautious monotony axiom, and it expresses as well a kind of skeptical property: come coherence is required of the conjunction of tokens entailed by the same information. The axiom says precisely that if

$$X \mid\!\sim a_1,\ X \mid\!\sim a_2,\ \cdots,\ X \mid\!\sim a_n,$$

then the set $X \cup \{a_1, a_2, \cdots, n\}$ must be coherent. (Note that this is much weaker than monotonicity, which says that if $X \mid\!\sim a$ and $Y \supseteq X$, then $Y \mid\!\sim a$.)

Is Axiom 3 a bit too strong? Should we never allow instances such as $X \mid\!\sim a$ and $X \mid\!\sim \neg a$ to hold at the same time? Although one can argue for not taking Axiom 3, it all depends on the kind of nonmonotonic entailment to be captured. Here, our skeptical version of entailment makes it reasonable to require Axiom 3; this will be confirmed later by our representation theorem. It is important to note that nonmonotonic entailment is different from default rules, where one does allow

$$\frac{X : a}{a},\ \ \frac{X : \neg a}{\neg a}$$

to appear in the same default set. However, defaults are lower level objects compared to nonmonotonic entailment. They are primitive construction rules for building models, and do not express any entailments on their own.

Axiom 5 is the *cautious cut*. We call it 'cautious cut' because sometimes the cut axiom takes the following form:

$$X \mid\!\sim T \ \&\ T, Y \mid\!\sim b \Rightarrow X, Y \mid\!\sim b.$$

This axiom is equivalent to the more restricted version

$$X \mid\!\sim T \ \&\ T, X \mid\!\sim b \Rightarrow X \mid\!\sim b$$

in the monotonic case, but definitely not in the nonmonotonic case.

Axiom 5 departs most from the corresponding axiom for information systems. By the fifth axiom for an information system, together with the other axioms, one can derive the following monotonicity property:

$$[X \vdash a \ \&\ (X \supseteq Y \in \mathit{Con})] \Rightarrow Y \vdash a.$$

But we do not even require cautious monotony to hold for a nonmonotonic system.

Given a default information structure $(A, \mathit{Con}, \vdash, \Delta)$, what is the appropriate entailment relation associated with it? Would it be reasonable to let $X \mid\!\sim a$ if $\dfrac{X : a}{a}$ is a default rule? Although this may seem to be a reasonable translation at first glance, it turns out to be too simple-minded. More often than not, no extra conditions are imposed on defaults, as in the case of the conflicting default rules above. Therefore, we rule out this definition.

The intuitive interpretation of a nonmonotonic entailment $X \mathrel{|\!\sim} a$ in default model theory is that, if X is our current state of knowledge, then a is believed. If we take extensions as plausible worlds, then $X \mathrel{|\!\sim} a$ can also be interpreted as a is true in all plausible superworlds of (the monotonic closure) of X. We point out that this interpretation is a special case of the general definition we introduced in Section 1.3. We mentioned in that section that

> With respect to a given default structure, a formula φ nonmonotonically entails a formula ψ, written as
> $$\varphi \mathrel{|\!\sim} \psi,$$
> iff ψ is satisfied in every extension of any minimal partial model of φ.

In the case that φ is a finite consistent set of tokens X, the minimal partial model for X is just $\overline{X}$, the deductive closure of X. We arrive at the following definition.

Definition 4.2 *Let $(A, Con, \vdash, \Delta)$ be a normal default structure. We define*

$$X \mathrel{|\!\sim}_A a$$

if

$$\forall y[\overline{X}\epsilon_A y \Rightarrow a \in y],$$

where ϵ_A is the extension relation for $\underline{A}$.

For cases where extensions exist (which is true for normal default structures), $X \mathrel{|\!\sim} a$ is equivalent to

$$a \in \bigcap\{y \mid \overline{X}\epsilon y\},$$

where subscripts are omitted.

First observe that the nonmonotonic entailment relation determined by a default structure does not in general have the cautious monotony property.

Example. Consider the following normal default structure $(A, Con, \vdash, \Delta)$, where

$$\begin{array}{l} A = \{a, b, c\}, \\ \Delta = \{\frac{: b}{b}, \frac{b : a}{a}, \ \frac{a : c}{c}\}, \\ \{a, b, c\} \notin Con. \end{array}$$

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

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

We have, therefore, $\emptyset \mathrel{|\!\sim} a$, $\emptyset \mathrel{|\!\sim} b$, but $\{a\} \not\mathrel{|\!\sim} b$.

□

A more convincing example shows that we can have $\emptyset \mathrel{|\!\sim} a$, $\emptyset \mathrel{|\!\sim} b$, but neither $\{a\} \mathrel{|\!\sim} b$, nor $\{b\} \mathrel{|\!\sim} a$.

Example. Consider the normal default structure $(A, Con, \vdash, \Delta)$, where

$$\begin{array}{l}A = \{a, b, a_1, b_1, \overline{a_1}, \overline{b_1}\},\\ \Delta = \{\dfrac{: a_1}{a_1}, \dfrac{a_1 : a}{a}, \dfrac{: b_1}{b_1}, \dfrac{b_1 : b}{b}, \dfrac{a : \overline{a_1}}{\overline{a_1}}, \dfrac{b : \overline{b_1}}{\overline{b_1}}\},\\ \{a_1, \overline{a_1}\} \notin Con, \{a_1, \overline{b_1}\} \notin Con, \{b_1, \overline{b_1}\} \notin Con, \{b_1, \overline{a_1}\} \notin Con, \{b, \overline{a_1}\} \notin Con, \{a, \overline{b_1}\} \notin Con.\end{array}$$

There is a unique extension for $\emptyset$: $\{a, b, a_1, b_1\}$. However, $\{a, \overline{a_1}\}$ is an extension for $\{a\}$, and $\{b, \overline{b_1}\}$ is an extension for $\{b\}$. Therefore, $\emptyset \mathrel{|\!\sim} b$, $\emptyset \mathrel{|\!\sim} a$, but neither $b \mathrel{|\!\sim} a$, nor $a \mathrel{|\!\sim} b$.

$\square$

What we are going to show next is a general result: the $\mathrel{|\!\sim}$ relation associated with every default information structure forms a nonmonotonic system. The following lemma will be needed to show the cautious cut axiom holds.

Lemma 4.1 *If $\overline{Z}\epsilon t$ and $P \subseteq t$, then $\overline{Z \cup P}\epsilon t$.*

Proof. Recall that $\overline{Z}\epsilon t$ means

$$t = \bigcup_{i \in \omega} \phi(\overline{Z}, t, i).$$

We prove by mathematical induction that

$$t = \bigcup_{i \in \omega} \phi(\overline{Z \cup P}, t, i).$$

$\subseteq$: Clearly $\phi(\overline{Z \cup P},\ t,\ 0) \subseteq t$ by the assumption $P \subseteq t$. Suppose $\phi(\overline{Z \cup P},\ t,\ i) \subseteq t$. Then

$$\begin{array}{l}\phi(\overline{Z \cup P},\ t,\ i+1)\\ = \overline{\phi(\overline{Z \cup P},\ t,\ i) \cup \{a \mid \dfrac{X : Y}{a} \in \Delta \;\&\; X \subseteq \phi(\overline{Z \cup P},\ t,\ i) \;\&\; Y \cup t \in Con\}}\\ \\ \subseteq \overline{\phi(\overline{Z \cup P},\ t,\ i) \cup \{a \mid \dfrac{X : Y}{a} \in \Delta \;\&\; X \subseteq t \;\&\; Y \cup t \in Con\}}\\ \subseteq t.\end{array}$$

Therefore, $\phi(\overline{Z \cup P},\ t,\ i) \subseteq t$ for every $i \geq 0$.

$\supseteq$: Obviously $\phi(\overline{Z \cup P},\ t,\ 0) \supseteq \phi(\overline{Z},\ t,\ 0)$. Assume $\phi(\overline{Z \cup P},\ t,\ i) \supseteq \phi(\overline{Z},\ t,\ i)$. Then

$$\begin{array}{l}\phi(\overline{Z \cup P},\ t,\ i+1)\\ = \overline{\phi(\overline{Z \cup P},\ t,\ i) \cup \{a \mid \dfrac{X : Y}{a} \in \Delta \;\&\; X \subseteq \phi(\overline{Z \cup P},\ t,\ i) \;\&\; Y \cup t \in Con\}}\\ \\ \supseteq \overline{\phi(\overline{Z},\ t,\ i) \cup \{a \mid \dfrac{X : Y}{a} \in \Delta \;\&\; X \subseteq \phi(\overline{Z},\ t,\ i) \;\&\; Y \cup t \in Con\}}\\ = \phi(\overline{Z},\ t,\ i+1).\end{array}$$

Therefore $\phi(\overline{Z \cup P},\ t,\ i) \supseteq \phi(\overline{Z},\ t,\ i)$ for every $i \geq 0$.

$\square$

Theorem 4.1 *Let $(A, Con, \vdash, \Delta)$ be a normal default structure. Define the triple $(A, Con, \mid\!\sim_A$ $)$, with $X \mid\!\sim_A a$ iff $X \in Con$ and*

$$\forall y.[\overline{X} \epsilon y \Rightarrow a \in y].$$

Then $(A, Con, \mid\!\sim_A)$ is a nonmonotonic system.

Proof . The nontrivial part is the cautious cut. Let $X \mid\!\sim_A T,\ T, X \mid\!\sim_A a$. This means

$$T \subseteq \bigcap\{y \mid \overline{X} \epsilon_A y\}$$

and

$$a \in \bigcap\{y \mid \overline{X \cup T} \epsilon_A y\}.$$

Let t be an extension of $\overline{X}$ with $T \subseteq t$. By Lemma 4.1, t is an extension of $\overline{X \cup T}$. However, a belongs to every extension of $\overline{X \cup T}$. Therefore, $a \in t$ and

$$a \in \bigcap\{y \mid \overline{X} \epsilon_A y\}.$$

$\square$

Now we come to the main result of the paper — the converse of Theorem 4.1. *Every* nonmonotonic system is determined by some normal default structure.

When engaged in the preliminary work on this paper, we thought that attempting the converse of Theorem 4.1 was a bold goal, for several reasons.

- Although the idea of deriving a nonmonotonic entailment from defaults has appeared in the literature here and there, no one has studied the converse – deriving a set of defaults Δ from a nonmonotonic entailment relation $\mid\!\sim$, with the property that the nonmonotonic entailment $\mid\!\sim_\Delta$ determined by the default set Δ is the same as the original one ($\mid\!\sim\, = \,\mid\!\sim_\Delta$).
- A simple minded approach, which translate each instance of nonmonotonic entailment $X \mid\!\sim a$ into a default rule $\dfrac{X : a}{a}$, does not work. For example, suppose the nonmonotonic entailment contains two instances: $\emptyset \mid\!\sim a$, $\{b\} \mid\!\sim c$, but *not* $\{b\} \mid\!\sim a$. The simple minded approach would result in a default set

  $$\Delta = \{\frac{\emptyset : a}{a},\ \frac{\{b\} : c}{c}\}.$$

  However, the nonmonotonic entailment determined by the default set would include the instance $\{b\} \mid\!\sim_\Delta a$, which is not in the original nonmonotonic entailment relation.

- In fact, *the exact converse of Theorem 4.1 is not true.* Suppose we have a nonmonotonic system with $\emptyset \mid\!\sim a$, $\emptyset \mid\!\sim b$, but $\{a\} \not\mid\!\sim b$. There are no default structures with exactly the tokens a and b, which determine the nonmonotonic system. This is because without new tokens, the default system will have at least two defaults

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

  and a, b are consistent. However, that also means $\{a\} \mid\!\sim b$ in the default structure. Although this example seems to suggest the use of cautious monotony, we have already pointed out that the nonmonotonic entailment relation determined by a default information structure does not in general have this property.

- The best thing one can achieve is to find a default structure so that the nonmonotonic system can be faithfully embedded into the one determined by the default structure. New tokens are introduced for this purpose. The idea is for each instance $X \mid\!\sim a$, one introduces a new token (X, a), distinct from the existing ones. It remains to specify the roles the new tokens will play; in particular, how the consistency predicate is going to be extended to them. This turns out to be a fairly complicated task: the reader is encouraged to think about the problem prior to reading the proof below.

- As a corollary, the converse of Theorem 4.1 implies that all cautious monotonic relations can be faithfully represented by default structures. But default structures are more expressive than that. They can faithfully represent *any* reasonable nonmonotonic consequence relation, and we know *exactly* what they can represent.

Here is the 'converse' of Theorem 4.1.

Theorem 4.2 *Let $(A, Con, \mid\!\sim)$ be a nonmonotonic system. There is a normal default structure*

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

with $B \supseteq A$, and

$$X \in Con \ \textit{ iff } X \subseteq A \ \& \ X \in Con',$$

which determines the nonmonotonic entailment $\mid\!\sim$, i.e. for $X \subseteq A$ and $a \in A$,

$$X \mid\!\sim a \ \textit{ iff } \ X \mid\!\sim_B a.$$

Example. An example will be helpful to illustrate the idea of the proof. We would like to construct a default structure $\underline{B}$, which determines the nonmonotonic entailment generated by

$$\emptyset \mid\!\sim a, \ \emptyset \mid\!\sim b,$$

but

$$\{a\} \not\mid\!\sim b, \{b\} \not\mid\!\sim a.$$

The defaults are

$$\frac{:(\emptyset,a)}{(\emptyset,a)}, \frac{(\emptyset,a):a}{a},$$
$$\frac{:(\emptyset,b)}{(\emptyset,b)}, \frac{(\emptyset,b):b}{b},$$
$$\frac{a:(\{a\},a)}{(\{a\},a)}, \frac{b:(\{b\},b)}{(\{b\},b)},$$
$$\frac{\{a,b\}:(\{a,b\},a)}{(\{a,b\},a)}, \frac{\{a,b\}:(\{a,b\},b)}{(\{a,b\},b)}.$$

For the consistency predicate, we let

$$\begin{array}{l} \{(\emptyset,a),(\{a\},a)\} \notin Con, \{(\emptyset,b),(\{b\},b)\} \notin Con, \\ \{(\{a\},a),(\{b\},b)\} \notin Con, \\ \{(\{a,b\},t),(Y,s)\} \notin Con \ \text{ if } \ Y \neq \{a,b\}, \\ \{b,(\{a\},a)\}, \{a,(\{b\},b)\} \text{ are also inconsistent.} \end{array}$$

It is routine to check that $\mid\!\sim_B a$, $\mid\!\sim_B b$, but neither $b \mid\!\sim_B a$, nor $a \mid\!\sim_B b$.

□

We now give a uniform procedure to construct the required default structure

$$\underline{B} = (B, \vdash_B, Con_B, \Delta)$$

from a nonmonotonic entailment $(A, Con, \mid\!\sim)$.

The token set B is

$$A \cup \{(X,a) \mid X \mid\!\sim a\},$$

where (X,a) are distinguished new tokens. The idea is to introduce a new token for each instance of the entailment.

Let $\vdash_B$ be flat, i.e., $X \vdash_B a$ iff $a \in X$. This means $\vdash_B$ does not play a key role here; it is the trivial entailment.

The default set Δ is

$$\bigcup_{(Y,b)\in B} \{\frac{Y:(Y,b)}{(Y,b)}, \frac{\{(Y,b)\}:b}{b}\}.$$

That is, each instance of nonmonotonic entailment $Y \mid\!\sim b$ will induce two default rules: one is $\dfrac{Y:(Y,b)}{(Y,b)}$, and the other is $\dfrac{\{(Y,b)\}:b}{b}$. Note that new tokens are used in both rules.

The consistency predicate Con_B plays an important role in ensuring the desired effect on defaults. It is specified by extending the consistency predicate Con to new tokens in the following way.

1. For $(X,a),(Y,b) \in B$,

$$\{(X,a),(Y,b)\} \in Con_B \Leftrightarrow X = Y.$$

   This means two new tokens are consistent just in case they are talking about what can be nonmonotonically entailed by the same set.

2. For $(Y,b) \in B$,
$$X \cup \{(Y,b)\} \in Con_B \Leftrightarrow X \cap A \subseteq \widetilde{Y}$$
where
$$\widetilde{Y} := \{c \mid Y \mathrel{|\!\!\sim} c\}.$$
By Axiom 3, $\widetilde{Y}$ is consistent. This specification means that when old and new tokens are mixed, a set is consistent if and only if all the old tokens are nonmonotonic consequences of the same set the new tokens are talking about. Note that for a set of old tokens, consistency remains the same.

We show first that the construction gives us a default structure. We check that the consistency predicate defined in this way has the required properties. First, each individual token is indeed consistent. Suppose $Z \in Con_B$ and $W \subseteq Z$. Items (1) and (2) above can be checked and we conclude that $W \in Con_B$.

Next, it has to be shown that the default structure $\underline{B}$ has the required properties. We will show this using several lemmas.

Lemma 4.2 *Let $(A, Con, \mathrel{|\!\!\sim})$ be a nonmonotonic system. We have, for any consistent set X,*
$$\bigcap\{\widetilde{Y} \mid Y \subseteq X \subseteq \widetilde{Y}\} = \widetilde{X}.$$

Proof . $\subseteq$: We clearly have $X \subseteq X \subseteq \widetilde{X}$. The required inclusion easily follows. $\supseteq$: Suppose $Y \subseteq X \subseteq \widetilde{Y}$. Because $X \subseteq \widetilde{Y}$, we have $Y \mathrel{|\!\!\sim} X - Y$. If $X \mathrel{|\!\!\sim} a$, we can rewrite this as $X - Y, Y \mathrel{|\!\!\sim} a$. Now, applying cut, we get $Y \mathrel{|\!\!\sim} a$. Therefore $\widetilde{X} \subseteq \widetilde{Y}$.

□

Lemma 4.3 *Let W be a consistent set in the nonmonotonic system. Put*
$$\rho(W) = \widetilde{W} \cup \{(W,b) \mid W \mathrel{|\!\!\sim} b\}.$$
Then $\rho(W)$ is consistent in the default structure $\underline{B}$.

Proof . This is straightforward from the specification of the consistency predicate Con_B.
□

Lemma 4.4 *Given $X \in Con$, suppose that X' is such that $X' \subseteq X \subseteq \widetilde{X'}$. Then $\rho(X')$ is an extension of X in the default structure $\underline{B}$.*

Proof . Since $\vdash_B$ is flat, any consistent set, in particular $\rho(X')$, is an ideal element. To show $\rho(X')$ is an extension of X, it is enough to check that
$$\rho(X') = X \cup \{t \mid \frac{Y : t}{t} \in \Delta \ \&\ Y \subseteq \rho(X') \ \&\ \rho(X') \cup \{t\} \in Con_B\}.$$

Since

$$\frac{X:(X,b)}{(X,b)}, \frac{\{(X,b)\}:b}{b}$$

are in the default set Δ for each b such that $X \mid\!\sim b$, $\rho(X')$ clearly is a subset of the right hand side.

On the other hand, suppose

$$\frac{Y:t}{t} \in \Delta \ \&\ Y \subseteq \rho(X') \ \&\ \rho(X') \cup \{t\} \in Con_B.$$

If $t = (Z, b)$ for some $b \in A$, then by item 2 for Con_B, $X' = Z$. Thus t is already in $\rho(X')$. If $t = b$ for some $b \in A$, then from the style of defaults we know that $Y = \{(Z, b)\}$ for some Z, and this means $Z = X'$ which again implies $t \in \rho(X')$.

□

Lemma 4.5 *Fix $X \in Con$. Every extension of X is of the form $\rho(X')$ with $X' \subseteq X \subseteq \widetilde{X'}$.*

Proof . Suppose y is an extension of X. By definition,

$$y = \bigcup_{i\in\omega} \phi(X, y, i),$$

where $\phi(X, y, 0) = X$, and

$$\phi(X, y, i+1) = \phi(X, y, i) \cup \{b \mid \frac{Y:b}{b} \in \Delta \ \&\ Y \subseteq \phi(X, y, i) \ \&\ y \cup \{b\} \in Con_B\}.$$

The default rules are designed in such a way that if $b \in A \cap \phi(X, y, i)$ but $b \notin X$, then there must be some (X', b) already in $\phi(X, y, i)$, with the property that $X' \subseteq X$, and $X \subseteq \widetilde{X'}$. The latter is required by the consistency condition.

If no (X', b)s are in $\phi(X, y, 1)$, then $\phi(X, y, 1) \subseteq X$, and we consider $\phi(X, y, 2)$, and so on. Eventually, $(X', b) \in \phi(X, y, k)$ for the first k, with the properties $X' \subseteq X$ (implied by the applicability of the default rule) and $X \subseteq \widetilde{X'}$ (implied by the consistency requirement). For each $i > k$, we have

$$\phi(X, y, i) \subseteq \widetilde{X'} \cup \{(X', b) \mid (X', b) \in B\},$$

again by the consistency requirement. Since $y \cup \{(X', b)\}$ is consistent for some $b \in A$,

$$y \cup \{(X', c) \mid (X', c) \in B\}$$

is also consistent, and, moreover, $y \cap A \subseteq \widetilde{X'}$. This means we have $y = \rho(X')$.

□

Proof of Theorem 4.2. Let X, a be in A such that $X \mid\!\sim a$. Lemmas 4.4 and 4.5 say that the $\rho(X')$, with $X' \subseteq X \subseteq \widetilde{X'}$ are exactly the extensions of X. However, Lemma 4.2 implies that

$$A \cap \bigcap\{\rho(X') \mid X' \subseteq X \subseteq \widetilde{X'}\} = \widetilde{X}.$$

We have, therefore, $X \mid\!\sim_B a$. On the other hand, if $X \not\mid\!\sim a$, then $X \not\mid\!\sim_B a$ because the propositions from A which belong to all the extensions of X are exactly $\widetilde{X}$, the set of nonmonotonic consequences of X. This proves Theorem 4.2.

□

It is worth noting that the construction used in the proof of Theorem 4.2 also tells us that there is a one one correspondence between those subsets X' of X such that $X \subseteq \widetilde{X'}$, and extensions for X in the default structure $\underline{B}$.

5 Cautious Monotonic Systems

This section treats the cautious monotony axiom in default structures. We show that precondition free default structures give rise to nonmonotonic entailment relations satisfying this axiom. We also remark that uniqueness of extensions implies cautious monotony (the converse is not true). To better present the material, we give a name to the collection of nonmonotonic systems satisfying cautious monotony: *cautious monotonic systems.*

Definition 5.1 *A cautious monotonic system is a triple*

$$(A, Con, \mid\!\sim)$$

where $(A, Con, \mid\!\sim)$ is a nonmonotonic system which satisfies the additional axiom of cautious monotony:

$$6.\ X \mid\!\sim a \ \&\ X \mid\!\sim b \Rightarrow X, a \mid\!\sim b.$$

Our first result in this section is the observation that precondition free structures give rise to an extension relation supporting cautious monotony. We now consider only precondition free structures.

For the next lemma, some terminology will be useful: given a precondition free default structure $\underline{A} = (A, Con, \vdash, \Delta)$, the set of default conclusions of $\underline{A}$ is the set $\{a \mid \frac{:a}{a} \in \Delta\}$. Further, we say that a set B is compatible with a set x if $x \cup B \in Con$.

Lemma 5.1 *Let $(A, Con, \vdash, \Delta)$ be a precondition free default structure. Then $y \in |A|$ is an extension of $x \in |A|$ if and only if there is a subset B of the default conclusions of $\underline{A}$ which is (i) maximal with the property that x is compatible with B, and (ii) $y = \overline{x \cup B}$.*

Proof. The proof is straightforward from definition.

□

Lemma 5.1 is the key to the following theorem.

Theorem 5.1 *Suppose $(A, Con, \vdash, \Delta)$ is a precondition free default structure. Define the triple $(A, Con, \mid\!\sim_A)$, with $X \mid\!\sim_A a$ iff $X \in Con$ and*

$$\forall y.[\overline{X} \epsilon y \Rightarrow a \in y].$$

Then $(A, Con, \mid\!\sim_A)$ is a cautious monotonic system.

Proof . We need only verify that cautious monotony is satisfied because Theorem 4.1 concludes that $(A, Con, \mid\!\sim_A)$ is a nonmonotonic system. Let $X \mid\!\sim_A T$ and $X \mid\!\sim_A b$. We want $T, X \mid\!\sim_A b$. Suppose that y is an extension of (the monotonic closure of) $T \cup X$. Then there is a maximal set B of default conclusions compatible with $T \cup X$, such that $y = \overline{T \cup X \cup B}$. We want to show that $b \in y$. Define

$$z = \overline{X \cup B}.$$

We claim that z is an extension of $\overline{X}$. Clearly $\overline{X} \subseteq z$ and z is consistent. If z is not an extension of $\overline{X}$, it is because B is not maximal in the sense of Lemma 5.1. That is, there is some maximal C, a proper superset of B, compatible with X, and such that $w = \overline{X \cup C}$ is an extension of $\overline{X}$. By hypothesis, $T \subseteq w$, and we already have $X \subseteq w$. So C is a larger set of default conclusions than B, but compatible with $T \cup X$, violating the maximality of B. This contradiction proves that z is an extension of $\overline{X}$. Thus $b \in z$, and since $z \subseteq y$, we have $b \in y$ as desired.

□

It is worth noting the relation between uniqueness of extensions and the cumulative property. The following proposition says that if there is only one extension for x, then the extension can be constructed in a cumulative way, in the sense that it is enough to check consistency with the current situation (rather than with the yet-to-be found extension) when applying a default rule.

Proposition 5.1 *Let $(A, Con, \vdash, \Delta)$ be a normal default structure for which extensions are always unique. Then $y \in |A|$ is an extension of $x \in |A|$ if and only if*

$$y = \bigcup_{i \in \omega} \psi(x, i),$$

where

$$\begin{aligned}
&\psi(x, 0) = x,\\
&\psi(x, i+1) =\\
&\overline{\psi(x, i)} \cup \{a \mid \frac{X : a}{a} \in \Delta,\ X \subseteq \psi(x, i),\ \{a\} \cup \psi(x, i) \in Con\}.
\end{aligned}$$

Proof. The proof of Theorem 4.1 illustrates a canonical way to construct an extension. However, by choosing different maximal consistent subsets, one gets at different extensions. The uniqueness of extensions imply that there must be only one maximal subset to choose at each step. By Definition 4.4, for each $i \geq 0$, this maximal subset must be

$$\overline{\phi(x, y, i)} \cup \{a \mid \frac{X : a}{a} \in \Delta,\ X \subseteq \phi(x, y, i),\ \{a\} \cup \phi(x, y, i) \in Con\}.$$

This proves the proposition.

□

The following theorem specifies another class of normal default structures which give rise to cautious monotonic systems.

Theorem 5.2 *Suppose $(A, Con, \vdash, \Delta)$ is a normal default structure for which extensions are unique. Then the induced nonmonotonic entailment $\mid\!\sim_A$ satisfies the cautious monotony law.*

Proof. Apply Lemma 4.1 and the unique-extension property.

$\square$

What can we say towards categorizing cautious monotonic systems by default structures? First, it is easy to see that precondition free structures are not enough. Second, default structures with unique extensions are not enough either. Third, Theorem 4.2 implies that each cautious monotonic system can be represented by a normal default structure. We do not have at present a characterization of those normal default structures which induce cautious monotonic entailment relations under our interpretation of that entailment using extensions. We leave this as an open problem; but we remark later that by using *dilations* instead of extensions, we can give a representation of cautious monotonic systems exactly as in Theorem 4.2.

We also remark that if the axiom of cautious monotony is assumed, then the construction in Theorem 4.2 can be simplified slightly. We close the section with this construction.

Let's look at a concrete example first. We want to construct a default structure $\underline{B}$, which determines the nonmonotonic entailment generated by

$$\emptyset \mid\!\sim a,\ \emptyset \mid\!\sim b,$$

and

$$\{a\} \mid\!\sim b, \{b\} \mid\!\sim a.$$

The defaults are

$$\frac{:(\emptyset, a)}{(\emptyset, a)}, \frac{(\emptyset, a) : a}{a},$$
$$\frac{:(\emptyset, b)}{(\emptyset, b)}, \frac{(\emptyset, b) : b}{b}.$$

The consistency relation is trivial in the sense that everything is consistent.

It is routine to check that $\mid\!\sim_B b$, $\mid\!\sim_B a$, and further, $b \mid\!\sim_B a$, $a \mid\!\sim_B b$.

$\square$

The key difference from the general construction for Theorem 4.2 is that we do not need to introduce new tokens for trivial instances of nonmonotonic entailment such as $X \mid\!\sim a$, where $a \in X$.

In general, the default structure

$$\underline{B} = (B, \vdash_B, Con_B, \Delta)$$

is constructed from a cautious monotonic system $(A, Con, \mid\!\sim)$ as follows.

The token set B is

$$A \cup \{(X, a) \mid X \mid\!\sim a \ \&\ a \notin X\},$$

where (X, a) are distinguished new tokens. Note the difference from the construction in the previous section: here we require $a \notin X$ for (X, a) to be a token of B.

Let $\vdash_B$ be flat, i.e., $X \vdash_B a$ iff $a \in X$.

The default set Δ is

$$\bigcup_{(Y,b)\in B} \{\frac{Y : (Y, b)}{(Y, b)}, \frac{\{(Y, b)\} : b}{b}\};$$

almost the same as before, except that we do not allow default rules where b is already a member of Y.

The consistency predicate Con_B extends Con with the following clauses.

1. For $(X, a), (Y, b) \in B$,
$$\{(X, a), (Y, b)\} \in Con_B \Leftrightarrow X = Y.$$

2. For $(Y, b) \in B$,
$$X \cup \{(Y, b)\} \in Con_B \Leftrightarrow X \cap A \subseteq \widetilde{Y}.$$

This gives us a default structure $\underline{B}$ with the required properties. The proof follows the same structure as that of Theorem 4.2, with a couple of simplifications. Under the assumption of cautious monotony, we can deduce the stronger result $\widetilde{X} = \widetilde{Y}$ from $X \subseteq Y \subseteq \widetilde{X}$. The extensions of $\overline{X}$ are exactly sets of the form

$$\widetilde{X} \cup \{(Y, b) \mid (Y, b) \in B \ \&\ b \notin X\}$$

with $X \subseteq Y$ and $\widetilde{X} = \widetilde{Y}$.

6 Nonmonotonic Entailment in Scott Topology

This section introduces nonmonotonic entailment in a more general topological setting. Our purpose is twofold: one is to represent nonmonotonic entailment in a more abstract form so that ideas from domain theory may be directly applied. The other is to prove two new results characterizing conditions on defaults which ensure cautious monotony. These results are more illuminating in this general setting.

6.1 Abstract defaults and extensions

Recall (see Section 2) that a Scott domain $(D, \sqsubseteq)$ is a cpo which is consistently complete: every bounded set has a least upper bound. The set of compact elements of a cpo D is written as $\kappa(D)$.

Definition 6.1 *Let $(D, \sqsubseteq)$ be a Scott domain. A default set in D is a subset Λ of $\kappa(D) \times \kappa(D)$. We call a pair $(a, b) \in \Lambda$ a default and think of it as a rule $\frac{a : b}{b}$, though this is an abuse of notation.*

A notion of extension can be introduced on Scott domains.

Definition 6.2 *Let $(D, \sqsubseteq)$ be a Scott domain. Let Λ be a default set in D. Write $x\epsilon y$ for $x, y \in D$ if*

$$y = \bigsqcup_{i \geq 0} \phi(x, y, i),$$

where $\phi(x, y, 0) = x$, and

$$\phi(x, y, i+1) = \phi(x, y, i) \sqcup \bigsqcup\{b \mid (a, b) \in \Delta \;\&\; a \sqsubseteq \phi(x, y, i) \;\&\; b \uparrow y\}.$$

When $x\epsilon y$, we call y an (abstract) extension of x.

The following is a characterization of abstract extensions, where $\sqcup$ stands for least upper bound, and $\sqcap$ stands for greatest lower bound.

Theorem 6.1 *For a Scott domain $(D, \sqsubseteq)$ and a subset $\Lambda \subseteq \kappa(D) \times \kappa(D)$, we have $x\epsilon y$ if and only if*

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

Proof. We prove a stronger result: for any y,

$$\bigsqcup_{i \geq 0} \phi(x, y, i) = \bigsqcap\{t \mid t = x \sqcup \bigsqcup\{b \mid (a, b) \in \Lambda \;\&\; a \sqsubseteq t \;\&\; b \uparrow y\}\}.$$

We first show that

$$\bigsqcup_{i \geq 0} \phi(x, y, i) \sqsubseteq \bigsqcap\{t \mid t = x \sqcup \bigsqcup\{b \mid (a, b) \in \Lambda \;\&\; a \sqsubseteq t \;\&\; b \uparrow y\}\}.$$

This is done by mathematical induction on i, to show that whenever

$$t = x \sqcup \bigsqcup\{b \mid (a, b) \in \Lambda \;\&\; a \sqsubseteq t \;\&\; b \uparrow y\},$$

we have $\phi(x, y, i) \sqsubseteq t$ for all i. Clearly

$$\phi(x, y, 0) \sqsubseteq t.$$

Suppose

$$\phi(x, y, i) \sqsubseteq t.$$

It is enough to show that

$$\bigsqcup\{b \mid (a, b) \in \Lambda \;\&\; a \sqsubseteq \phi(x, y, i) \;\&\; b \uparrow y\}$$

$$\sqsubseteq x \sqcup \bigsqcup\{b \mid (a, b) \in \Lambda \;\&\; a \sqsubseteq t \;\&\; b \uparrow y\}.$$

But this is clear from the assumption that $\phi(x, y, i) \sqsubseteq t$.

We now show that

$$\bigsqcup_{i\geq 0}\phi(x,y,i) \sqsupseteq \bigsqcap\{t \mid t = x \sqcup \bigsqcup\{b \mid (a,b)\in\Lambda \;\&\; a \sqsubseteq t \;\&\; b\uparrow y\}\}.$$

This is done by demonstrating that $\bigsqcup_{i\geq 0}\phi(x,y,i)$ is one of the t's, that is,

$$\bigsqcup_{i\geq 0}\phi(x,y,i) = x \sqcup \bigsqcup\{b \mid (a,b)\in\Lambda \;\&\; a \sqsubseteq \bigsqcup_{i\geq 0}\phi(x,y,i) \;\&\; b\uparrow y\}.$$

However, the above follows from the fact that a's are isolated elements and

$$\{\phi(x,y,i) \mid i\geq 0\}$$

is an ω-increasing chain.

□

Theorem 6.1 suggests the use of fix-point theorems in domain theory (or lattice theory). Recall that a function $f : D \to D$ is continuous if it is monotonic and it preserves least upper bounds of ω increasing chains. The least fix point of any such continuous function can be constructed by taking the least upper bound of the ω chain

$$\perp \sqsubseteq f(\perp) \sqsubseteq f(f(\perp)) \sqsubseteq \cdots$$

For a fixed domain D and default Λ, let

$$\begin{aligned}&\xi(x,u,v) = x \sqcup \bigsqcup\{b \mid (a,b)\in\Lambda \;\&\; a\sqsubseteq u \;\&\; b\uparrow v\},\\ &\eta(x,v) = \bigsqcap\{\, t \mid t = \xi(x,t,v)\,\}.\end{aligned}$$

It is easy to check that for fixed x and v, $\xi(x,u,v)$ is a continuous function in u. Therefore, $\xi(x,u,v)$ has a least fix point, such that $\xi(x,l,v) = l$, which is exactly what $\eta(x,v)$ trying to express. Indeed, the proof of Theorem 6.1 also confirms that

$$\xi(x,\eta(x,v),v) = \eta(x,v).$$

We have the following representation theorem, which builds on top of Scott's representation theorem for information systems. The proof is straightforward, hence omitted.

Theorem 6.2 *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* $\frac{X:Y}{Y}$ *to the pair* (x,y) *of compact elements determined by* (X,Y)*; and conversely via the 'inverse' correspondence from Scott domains to information systems.*

Note that in this representation theorem a slightly more general form of defaults $\frac{X:Y}{Y}$ is used, where Y is a finite set instead of a singleton. However, all previous results directly generalize to this case.

6.2 Nonmonotonic entailment between open sets

Let $(D, \sqsubseteq)$ be a Scott domain. A subset $U \subseteq D$ is said to be *Scott open* if (i) U is upward closed: $x \in U$ and $x \sqsubseteq y$ imply $y \in U$; and (ii) for any ω-increasing chain

$$x_0 \sqsubseteq x_1 \sqsubseteq x_2 \sqsubseteq \cdots,$$

$\bigsqcup_i x_i \in U$ implies $x_k \in U$ for some k.

One checks readily that under this definition of 'open', that the collection of open subsets of a Scott domain 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. After a finite time, we find that the element does indeed have the property. Such properties are sometimes called 'affirmable' [24].

It is straightforward to prove the following in any Scott domain D.

Theorem 6.3 Compactness in the Scott topology

1. *For each finite element $x \in D$, the set $\uparrow x = \{u \mid x \sqsubseteq u\}$ is open. We call it* prime open.

2. *Every open set U is the union of the prime opens generated by the compact elements of U.*

3. *Every compact open set X is a finite union of such prime opens.* (Compact here means the topological usage: every covering of X by open sets has a finite subcovering.)

A set Λ of defaults on a Scott domain $(D, \sqsubseteq)$ induces a nonmonotonic entailment relation between open sets. For any open set U, let μU stands for the set of minimal elements of U. Because U is an open set, elements in μU are isolated (compact). Suppose U, V are open sets of D. With respect to Λ, define $U \mid\!\sim V$ iff given any element x in μU, every extension y of x is a member of V. This, of course, corresponds to the more concrete description given in Section 1.3. Unsurprisingly, cautious monotony fails in this general setting: from $U \mid\!\sim V$ and $U \mid\!\sim W$, we cannot conclude $U \cap V \mid\!\sim W$. This is because general open sets allow disjunctions, in the form of unions of prime opens. Our representation theorems are really about nonmonotonic entailment relations on prime opens. But because prime opens are determined as the upward closure of compact elements, we can also regard this entailment as a relation between compact elements. This is what we called the "abstract setting" for entailment relations in the introductory section. We reserve the notation $\mid\!\sim$ for entailment between open sets, and use $\rightsquigarrow$ as a relation on $\kappa(D)$. (Think of this as the entailment relation on compact opens.) If now Λ is a default set on D, we define the relation $\rightsquigarrow_\Lambda$ in analogy with our default entailment relation: we let $a \rightsquigarrow_\Lambda b$ iff for every extension e of a, we have $b \sqsubseteq e$.

Our representation theorem for nonmonotonic systems is now the following.

Theorem 6.4 *For any nonmonotonic system $(A, Con, \mid\!\sim)$ we can find a Scott domain D, a default set Λ, and an injective mapping $\theta : Con \to \kappa(D)$ satisfying the following: For any X and Y in Con, $X \mid\!\sim Y$ if and only if $\theta(X) \leadsto_\Lambda \theta(Y)$.*

This result is just a corollary of Theorem 4.2 and the representation of monotonic information systems as Scott domains.

A more interesting kind of possible representation theorem is based on some observations about the relation $\leadsto_\Lambda$. Observe that it satisfies the following laws (we suppress the subscript):

- **Reflexivity:** $a \leadsto a$ for all compact a;
- **Right Weakening:** if c is compact and $c \sqsubseteq a$ then $a \leadsto c$;
- **Consistency:** if $a \leadsto b$ then $a \uparrow b$;
- **Right Conjunction:** If F is a finite subset of $\kappa(D)$ and $a \leadsto b$ for all $b \in F$ then $a \leadsto \bigsqcup F$ (note that in particular F is consistent);
- **Cautious cut:** If $a \leadsto b$ and $a \sqcup b \leadsto c$ then $a \leadsto c$.

Notice that $\leadsto$ is somewhat like $\sqsupseteq$: it could be called "default subsumption". The Cautious Cut rule is the substitute for transitivity. It would seem possible to start with a Scott domain, and a consequence relation satisfying the above laws, and then represent it by means of abstract defaults in an embedding domain. This topic, though, belongs to a later paper, where we plan to discuss more issues in the abstract model theory of defaults.

6.3 Trace sets ensure cautious monotony

It has been mentioned in Section 5 that cautious monotony can be regained when extensions are unique, in the model theoretic setting. The 'model theoretic' setting in a topology corresponds to restricting the open sets to prime opens. However, the condition of 'unique extension' is not very useful, because it is not realistic to verify. In this subsection we provide a sufficient condition for default sets to determine unique extensions. This gives a concrete and efficient way to verify cautious monotony by checking a property of the default set.

Our result is motivated from results in domain theory. It is well-known in domain theory that a function $f : D \to D$, where D is a Scott domain, is continuous iff for every finite element $b \in \kappa(D)$, $b \sqsubseteq f(x)$ implies $b \sqsubseteq f(a)$ for some $a \sqsubseteq x$, with a finite. This means, in a sense, that continuous functions are determined by pairs of finite elements. But what kind of properties must a set of such pairs have to ensure that it correspond to a continuous function?

We have

Theorem 6.5 *A set $\{ (a_i, b_i) \mid i \in I \} \subseteq \kappa(D) \times \kappa(D)$ determines a continuous function if for all finite set $J \subseteq I$, $\{ a_i \mid i \in J \}$ bounded above implies $\{ b_i \mid i \in J \}$ is bounded above.*

The proof for Theorem 6.4 is quite routine, and we omit it. However, we point out that the function determined by the set $\{ (a_i, b_i) \mid i \in I \}$ is the following:

$$f(x) = \bigsqcup\{b_k \mid a_k \sqsubseteq x \& k \in I\}.$$

We call a set with the property mentioned in the previous theorem a *trace set.*

Definition 6.3 *Let $F \subseteq \kappa(D) \times \kappa(D)$, where D is a Scott domain. F is a trace set if for every finite subset G of F, the consistency of G's first components implies the consistency of G's second components, i.e.,*

$$\pi_1 G \uparrow \Rightarrow \pi_2 G \uparrow .$$

Of course, here π_i's are projections to the ith component, and $X \uparrow$ means that X is bounded above in D.

Let's define, for a function $f : D \to D$, the set $\mathsf{tr}(f)$ to be a set of pairs (a, b) in $\kappa(D) \times \kappa(D)$ such that $b \sqsubseteq f(a)$. Clearly $\mathsf{tr}(f)$ is a trace set, and we call it the trace set of f. The set $\mathsf{tr}(f)$ has the following additional property of *saturation*:

$$[(a, b) \in \mathsf{tr}(f) \ \& \ a \sqsubseteq a' \ \& \ b' \sqsubseteq b] \Rightarrow (a', b') \in \mathsf{tr}(f).$$

However, when constructing a function from a trace set, the 'derived' pairs like (a', b') does not contribute anything.

Observe that if we start with a continuous function $f : D \to D$, construct its trace set $\mathsf{tr}(f)$, and then derive a function from it via Theorem 6.4, we get back the same function. On the other hand, if we start with a *saturated* trace set, derive a continuous function from the set, and then find the trace set of the function, we get exactly the same trace set back.

We now come to the main theorem of this subsection.

Theorem 6.6 *Let Λ be an abstract default set of a Scott domain D. Then extensions are unique for Λ if*

$$\Lambda \cup I_{\kappa(D)}$$

is a trace set, where $I_{\kappa(D)} = \{(a, a) \mid a \in \kappa(D)\}$.

As a consequence of this theorem, one can easily check if the nonmonotonic consequence induced by a default structure $(A, Con, \vdash, \Delta)$ satisfies cautious monotony. One simply augments Δ with all trivial instances $\dfrac{\{a\} : a}{a}$, for $a \in A$, then check that for any finite set

$$\frac{X_1 : a_1}{a_1}, \ \frac{X_2 : a_2}{a_2}, \ \cdots, \ \frac{X_n : a_n}{a_n}$$

from Δ, if

$$X_1 \cup X_2 \cup \cdots \cup X_n \in Con,$$

then

$$\{a_1, a_2, \cdots, a_n\} \in Con.$$

Note that adding $\dfrac{\{a\} : a}{a}$ into the default set seems quite innocent, but with the trace set property, it produces some extra effect, as can be seen in the following proof.

Proof of Theorem 6.6. Suppose Λ has the property mentioned in Theorem 6.6. Clearly Λ itself is a trace set and, moreover, the function determined by $\Lambda \cup I_{\kappa(D)}$ is greater than or equal to the identity function under the pointwise order.

Given any element $x \in D$, we have the following monotonic procedure for building an extension, by taking advantage of the trace set property of $\Lambda \cup I_{\kappa(D)}$.

Let f denote the continuous function determined by $\Lambda \cup I_{\kappa(D)}$. Clearly f is *inflationary*, in the sense that $f(t) \sqsupseteq t$ for every t. There is a canonical way to construct a fix point for such an inflationary function: just take the least upper bound of

$$x \sqsubseteq f(x) \sqsubseteq f(f(x)) \sqsubseteq \cdots .$$

Important to us is the fact that *this fix point is an extension for x*. Moreover, it can be readily shown that it is indeed the unique extension of x.

□

Although Theorem 6.6 captures a large class of default sets which determine unique extensions, one wonders if the condition in Theorem 6.6 is also necessary for unique extensions. The answer is, unfortunately, no. It is not hard to find finite examples which confirm this: the reader is encouraged to find such examples.

To close the section, we briefly mention another condition for cautious monotony.

Theorem 6.7 *For abstract defaults on Scott domains, the derived entailment has the cautious monotony property if Λ satisfies the following condition:*

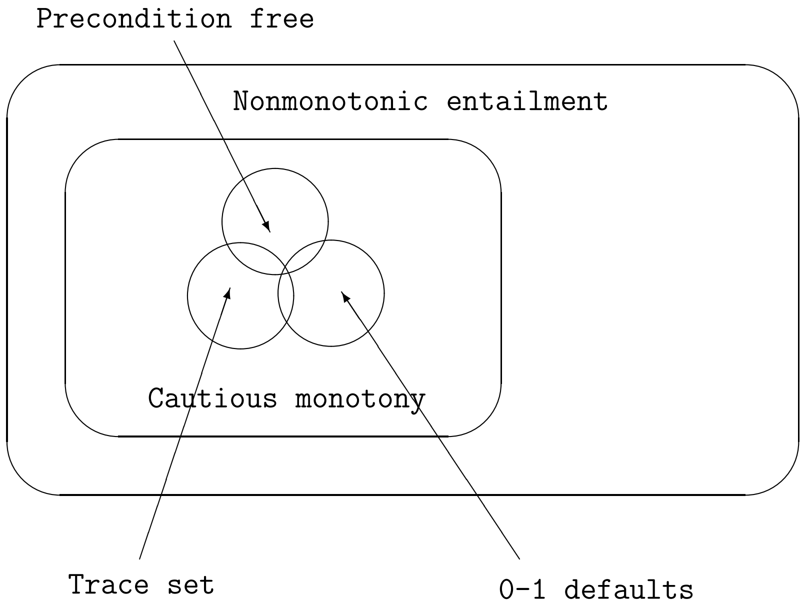
$$[(a, b), (a', b') \in \Lambda \;\&\; a \uparrow a'] \Rightarrow [b = b' \quad or \; b \not\uparrow b'].$$

Proof. We let $\rightsquigarrow_\Lambda$ be the relation on compact elements of D mentioned above, and we drop the subscript Λ. Then the Cautious Monotony property is that whenever $x \rightsquigarrow y$ and $x \rightsquigarrow z$, we have that $x \sqcup y \rightsquigarrow z$. Suppose $x \rightsquigarrow y$ and $x \rightsquigarrow z$. To show that $(x \sqcup y) \rightsquigarrow z$, let $x \epsilon t$. Applying the assumed property for Λ, we can see that $t = x \sqcup \bigsqcup B$, where B is either a singleton, or empty. If B is empty, then we have $x \sqsupseteq z$, which implies $x \sqcup y \sqsupseteq z$. If $B = \{b\}$, then we have $x \sqcup b \sqsupseteq z$. Therefore, any extension for $x \sqcup y$ is of the form $x \sqcup y \sqcup b$, which dominates z. This proves $(x \sqcup y) \rightsquigarrow z$.

□

For lack of a better name, let's call defaults with the property mentioned in the previous theorem *0-1 defaults*. It should be pointed out that for this case, cautious monotony holds not because of unique extensions, but because there are so many incompatible extensions.

The following picture illustrates the relationships among different classes of default structures.



7 Conclusion

This paper views the issues involved in the study of nonmonotonic entailment in the light of default model theory. It is an integral part of a more ambitious effort to reformulate nonmonotonic reasoning. We have proved a representation theorem which characterizes the properties of a nonmonotonic entailment relation derived from default rules and extensions. This establishes a close connection between defaults and the abstract nonmonotonic entailment relation. Although we have at least one convincing example to show that cautious monotony may not hold in general, subclasses of defaults are identified for which cautious monotony does hold, in the model theoretic setting. We then represent the abstract nonmonotonic entailment relation in a topological setting. This also makes it possible to apply results in domain theory. Although Section 5 and Section 6 present several subclasses of defaults for which the induced nonmonotonic entailment relation satisfies cautious monotony, the precise characterization of such defaults remains an open question. However, it is important to point out that properties of derived nonmonotonic entailment not only depend on the kind of defaults used, but also depend on the procedure for building the extended partial world. Although extension is one of the key model building method, there are other possibilities. In [21], we introduce a construction called a *dilation*, motivated from the need to ensure the existence of extended partial world for all reasonable defaults. Dilations are a robust generalization of extensions, and exist for all semi-normal default structures.

What happens if, instead of using extensions, we use dilations in the passage from a default structure to a nonmonotonic entailment relation? The answer is, quite unexpectedly, that these capture exactly the cumulative entailment relations: those satisfy cautious monotony. We have obtained representation results similar to those given in Section 4. This somehow makes the 'open question' mentioned in the previous paragraph less urgent than it might seem. Extensions seem to be fundamentally at odds with cautious monotony. Dilations, however, are in harmony with cumulative reasoning. We plan to report results on dilation-related nonmonotonic entailment in a forthcoming paper.

References

[1] Samson Abramsky and Achim Jung. Domain theory. In *Handbook of Logic in Computer Science, vol 3*, Oxford Science Publications, 1994 (to appear).

[2] E. Adams. *The logic of conditionals*. Netherlands:D. Reidel, 1975.

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

[4] Philippe Besnard. *An introduction to default logic*. Springer-Verlag, 1989.

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

[6] D. Gabbay. Theoretical foundations for nonmonotonic reasoning in expert systems. In K. R. Apt, editor, *Proceedings of NATO Advanced Study Institute on Logics and Models of Concurrent Systems*, pages 439–457. Springer Verlag, 1985.

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

[8] Carl Gunter. *Semantics of Programming Languages: Structures and Techniques*. The MIT Press, Boston, 1992.

[9] J. Halpern and Y. Moses. A guide to modal logics of knowledge and belief. In *Proc. IJCAI 85*, pages 480–490, 1985.

[10] J. Halpern and M. Vardi. Model checking vs. theorem proving – a manifesto. In *Proceedings of Second Annual Conference on Knowledge Representation*, pages 118–127. Morgan Kaufmann, 1991.

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

[12] D. Lehmann. What does a conditional knowledge base entail? In *Proceedings of KR '89*, pages 212–234. Morgan Kaufmann, 1989.

[13] H. J. Levesque. Making believers out of computers. *Artificial Intelligence*, 30:81–108, 1986.

[14] H. J. Levesque. All i know: A study in autoepistemic logic. *Artificial Intelligence*, 42:263–309, 1990.

[15] V. Lifschitz. On open defaults. In *Proceedings of the Symposium on Computational Logics*, Brussels, 1990.

[16] D. Makinson. General theory of cumulative inference. In *Proceedings of the second international conference on nonmonotonic reasoning, Lecture Notes in Artificial Intelligence 346*. Springer-Verlag, 1989.

[17] J. Pearl. From adams' conditionals to default expressions, causal conditionals, and counterfactuals. In *Festschrift for Ernest Adams*, Cambridge University Press, 1993, to appear.

[18] D. L. Poole. A logical framework for default reasoning. *Artificial Intelligence*, 36(1):27–48, 1988.

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

[20] William Rounds and Guo-Qiang Zhang. Logics for default domains. Submitted for publication, 1993.

[21] William Rounds and Guo-Qiang Zhang. Domain theory meets default logic. *International Journal of Logic and Computation*, 1994, to appear.

[22] William Rounds and Guo-Qiang Zhang. Logical considerations on default semantics. In *Presentation at the Third International Symposium on Artificial Intelligence and Mathematics, Fort Lauderdale, Florida*, January, 1994.

[23] Dana S. Scott. Domains for denotational semantics. In *Lecture Notes in Computer Science 140*, 1982.

[24] S. Vickers. *Topology via Logic*. Cambridge University Press, 1989.

[25] Glynn Winskel. *The Formal Semantics of Programming Languages*. The MIT Press, Boston, 1992.

[26] Guo-Qiang Zhang. *Logic of Domains*. Birkhauser, Boston, 1991.